Semiconductor Interface Specialists Conference SISC 2008
San Diego, CA, USA
December 10, 2008

Introduction to Charge Pumping and Its Applications

Guido Groeseneken
IMEC – KU Leuven

imec

Purpose

- To explain the basic principles of the charge pumping technique for characterising the interface charge in MOSFET’s
- To illustrate the application of the technique for the analysis of the degradation of MOSFET’s and MOS-related devices, for energy and spatial profiling of interface traps
- To discuss the effect of oxide thickness scaling and how Charge pumping can successfully be used for analysing high k dielectrics
- Targeting both novices as well as experts in the field

PURPOSE OF CHARGE PUMPING

IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. ED-16, NO. 3, MARCH 1969

Charge Pumping in MOS Devices

J. STEPHEN BRUGLER, MEMBER, IEEE, AND PAUL G. A. JESPERS, SENIOR MEMBER, IEEE

Abstract—Gate pulses applied to MOS transistors were found to stimulate a net flow of charge into the substrate. Investigation of this effect revealed a charge-pumping phenomeonon in MOS gate-controlled-diode structures. A first-order theory is given, whereby the injected charge is separated into two components. One component involves coupling via fast surface states at the Si–SiO_2 interface under the gate, while the other involves recombination of free inversion-layer charge into the substrate.

Introduction—The Basic Experiment

During the course of evaluating enhancement-type MOS transistors for a very-low-level switching application, a spurious signal having magnitude of a few percent of the expected capacitive transients was observed. The evidence suggested that a net charge was being injected across source– and drain-substrate junctions when gate pulses were applied.

The experiment of Fig. 1 clarified the effect. The substrate current of an MOS transistor was smoothed by a capacitor and fed to a dc ammeter. Source and drain were shorted and reverse biased ($V_R < 0$). In the absence of any gate pulses, the ammeter simply indicated the junctions' negative reverse leakage currents. When the *n*-type substrate was periodically inverted by negative gate pulses of width T_1, the current reversed polarity and became positive. Its magnitude increased with gate pulse frequency, becoming linear when leakage effects were swamped. If the reverse voltage was zero, the dc substrate current was linearly proportional to frequency over the limits of the measuring equipment. This is shown by Fig. 2 for the case of a 2N4066 transistor. This linearity is clearly indicative of a "charge pumping" action whereby a fixed charge is measured at each gate pulse. Since no dc component of the measured magnitude can flow through the oxide, this charge must be injected across the junctions. A current of 1.3 nA was measured at a 1-kHz frequency, so the charge delivered per pulse was 1.3 pC. Note that current is able to flow in the forward direction, opposite to the leakages even if the junctions are reverse biased. This means that power is being transferred from the pulse source to the battery.

The magnitude of the substrate current at a fixed frequency is plotted as a function of gate voltage in Fig. 3 for three values of junction bias. A threshold gate voltage is seen to exist, below which no current is

Manuscript received July 11, 1968; revised December 4, 1968. This work was supported by the Joint Electronics Program under Contract Nonr-225(83), NR 373 360.

J. S. Brugler is with Stanford University, Stanford, Calif.
P. G. A. Jespers is with the Catholic University of Louvain, Louvain, Belgium.

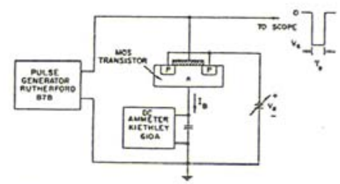


Fig. 1. Basic MOS charge pumping experiment.

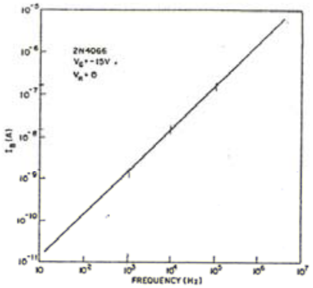


Fig. 2. Dc substrate current versus gate pulse frequency.

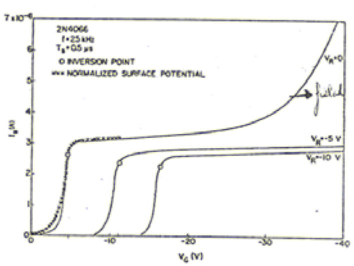


Fig. 3. Dc substrate current versus gate voltage.

measured. As the gate pulse amplitude increas[es, a] sharp rise in current is observed, followed by sa[tura]tion. The current remains saturated up to the d[…] voltage limits, except when $V_R = 0$, in which cas[e …] current plot curves upward at higher gate vol[t…]

Interface characterization techniques

	INDIRECT TECHNIQUES	DIRECT TECHNIQUES
MOS-Capacitors	• **LF-CV** • **HF-CV**	• Conductance (without gate leakage) • DLTS
MOS-Transistors	• Weak inversion • 1/f Noise	• Current DLTS • DCIV • **Charge pumping**

Classification of interface characterisation techniques

What can charge pumping do?

- Measure the interface state density at the Si-substrate/gate oxide interface
- Resolution <10^9 states/cm^2 or better, single trap capability
- Determine separate shifts in threshold and flatband voltages from interface trap generation
- Determine the energy distribution of interface states
- Give information on spatial position of interface states in the source-drain direction and/or in the depth direction
- Measure inversion charge density
- In high-k with thin interface layer : measure density of bulk high-k states

Outline

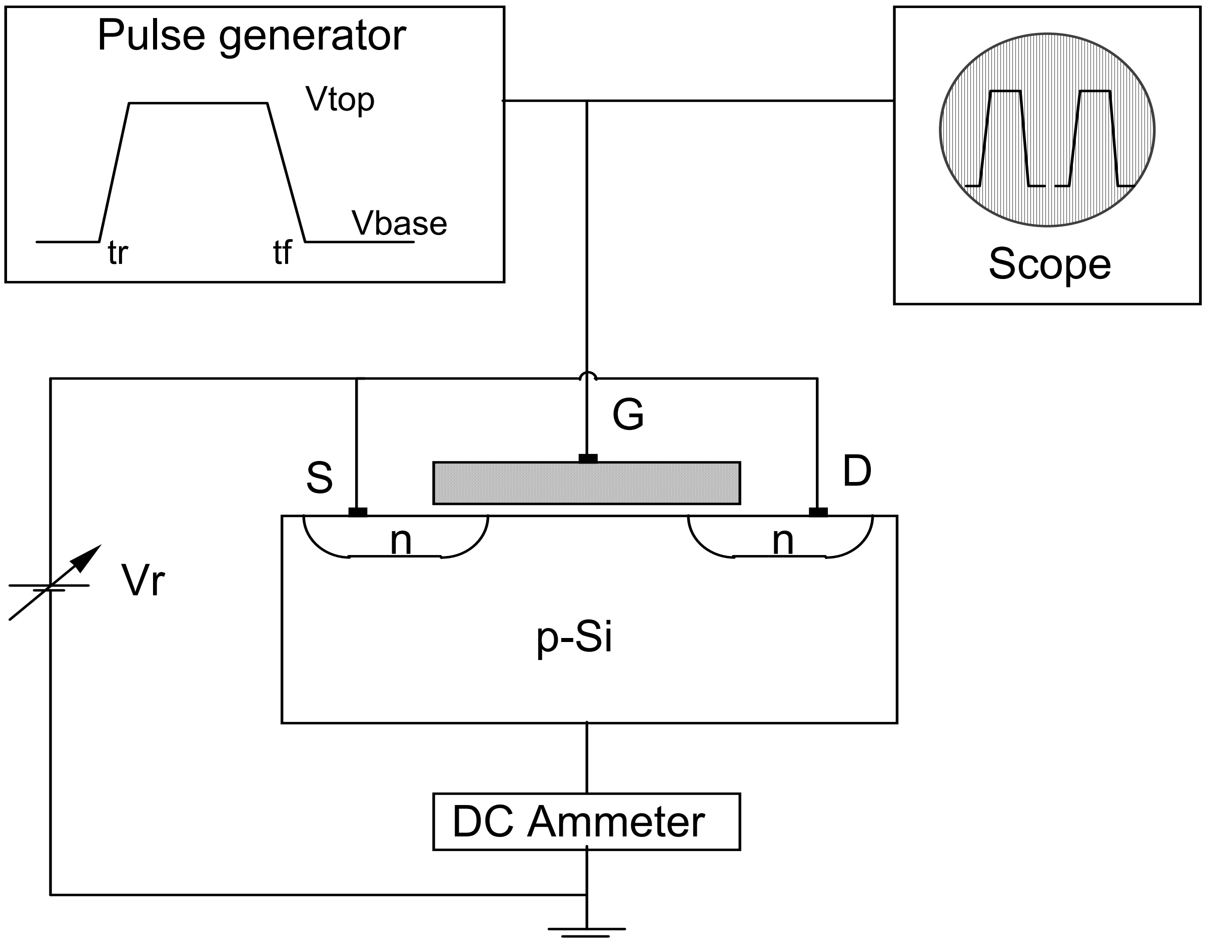
1. Introduction
2. Basic CP-principle
3. Second order model
4. Temperature dependence
5. Base level edges
6. MOSFET degradation
7. Energy distribution
8. Lateral and vertical profiling
9. Geometric components
10. Effects of oxide thickness scaling
11. Single trap characterization
12. CP in high k gate stacks

Limerick #1

What is charge pumping ?

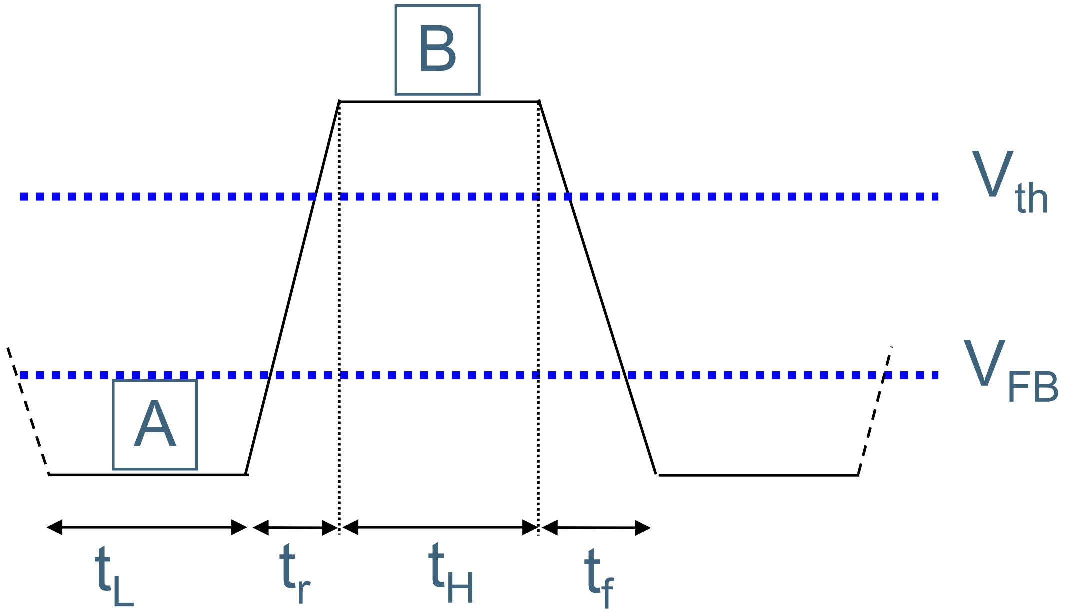
It's a method that works in 2 steps
First electrons are captured on traps
Then we pump in a hole
Causing current with the goal
To be proportional to the density of traps

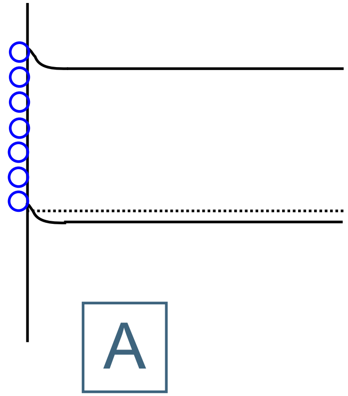
Basic principle of charge pumping

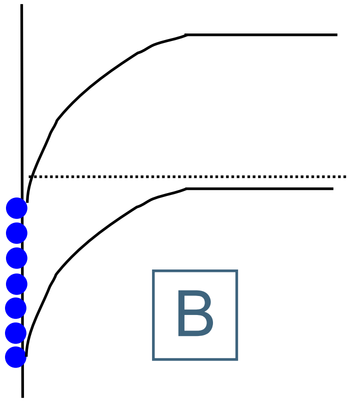


Experimental set-up: apply a gate pulse
measure the substrate current !

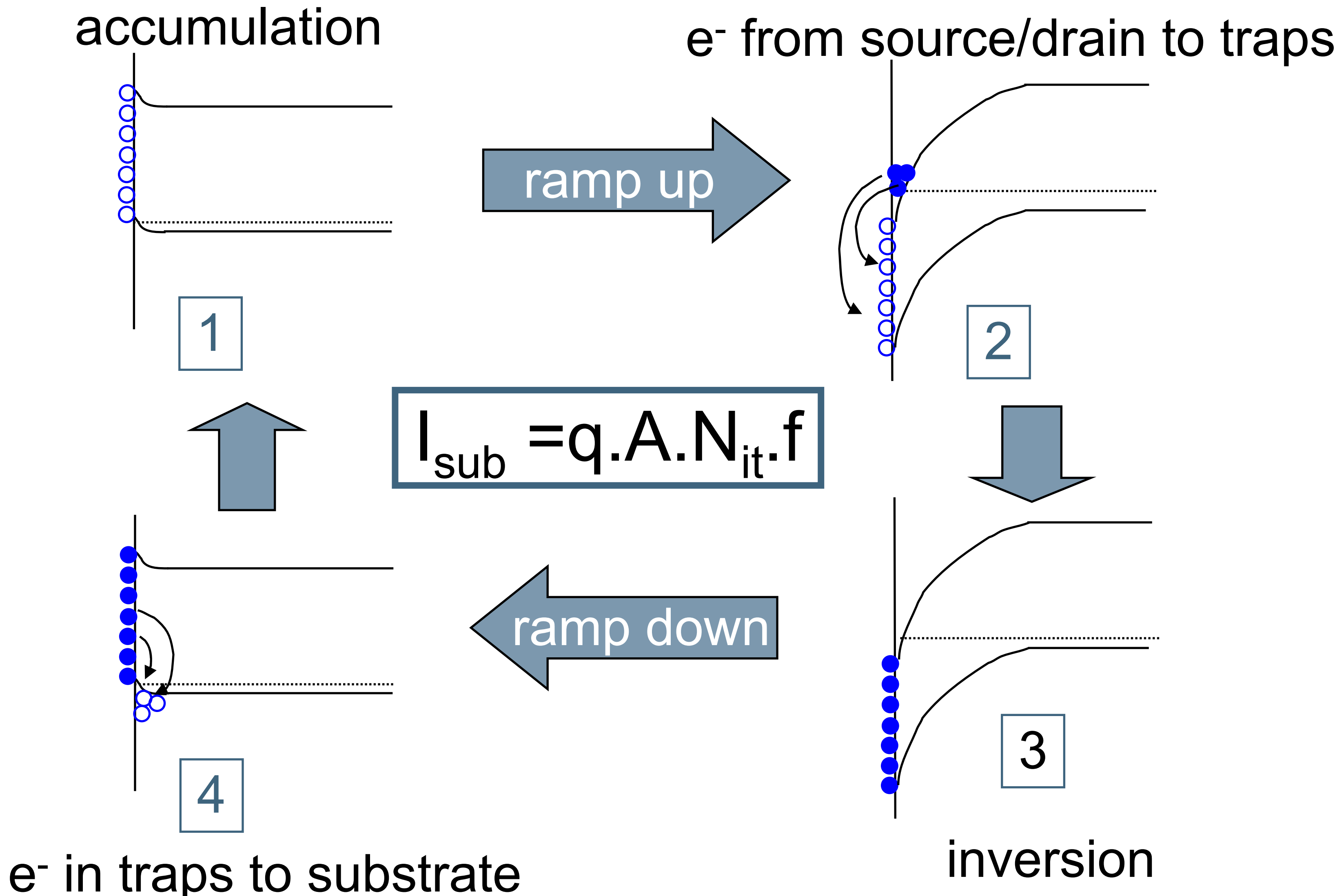
Sweeping between inverse and accumulation







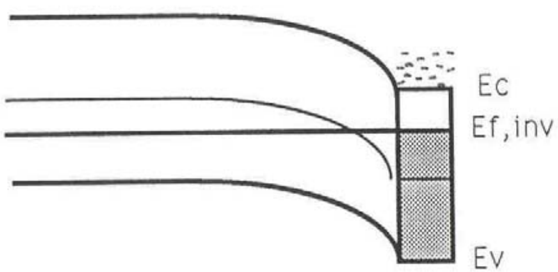
Charge pumping: 1st order understanding



$$I_{sub} = q.A.N_{it}.f$$

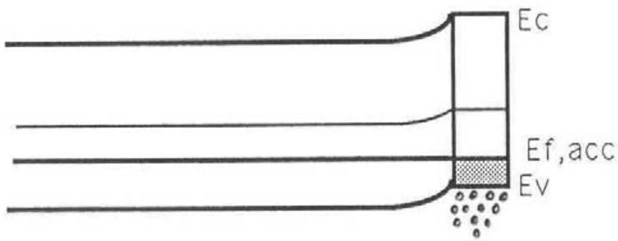
SIMPLE OPERATING CHARGE PUMPING PRINCIPLE

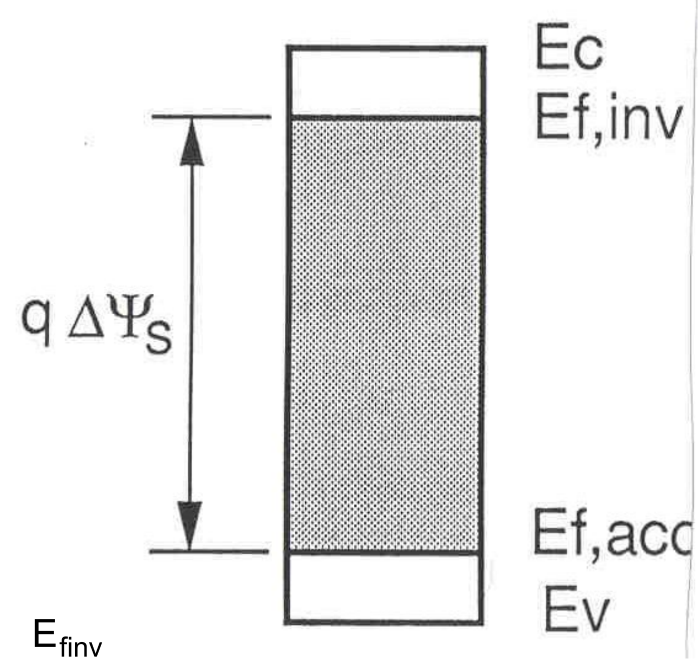
VG > VT



$$Q_{ss} = q\,A_G \int_{E_{facc}}^{E_{finv}} D_{it}(E)\,dE$$

VG < VFB

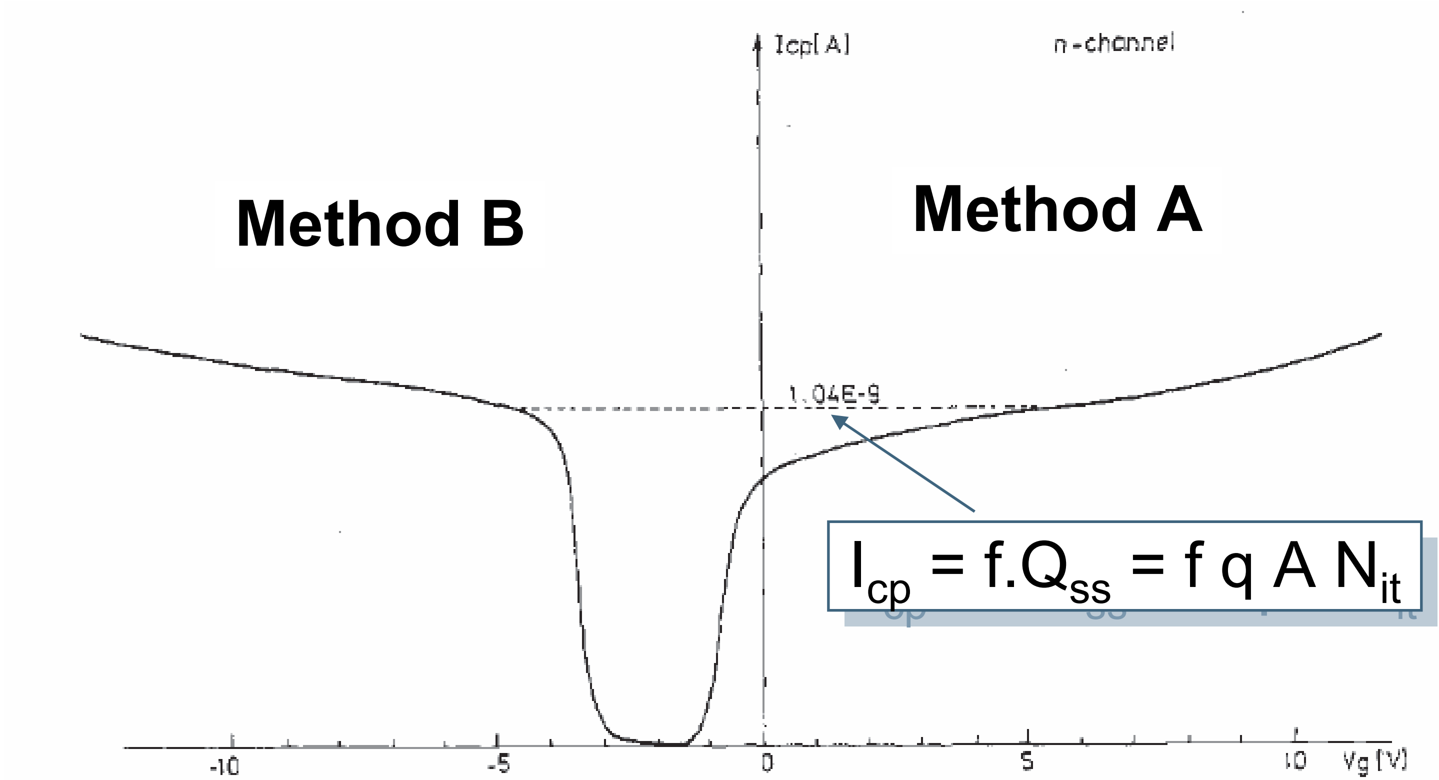




$$Q_{ss} = q\,A_G \int_{E_{facc}}^{E_{finv}} D_{it}(E)\,dE$$

$$\mathbf{I_{cp} = f.Q_{SS} = f\,q\,A_G\,\overline{D_{it}}\,\Delta\Psi_S}$$

Simple operating principle



Method B

Pulse level in inversion is fixed

Pulsing the surface into accumulation with increasing amplitudes

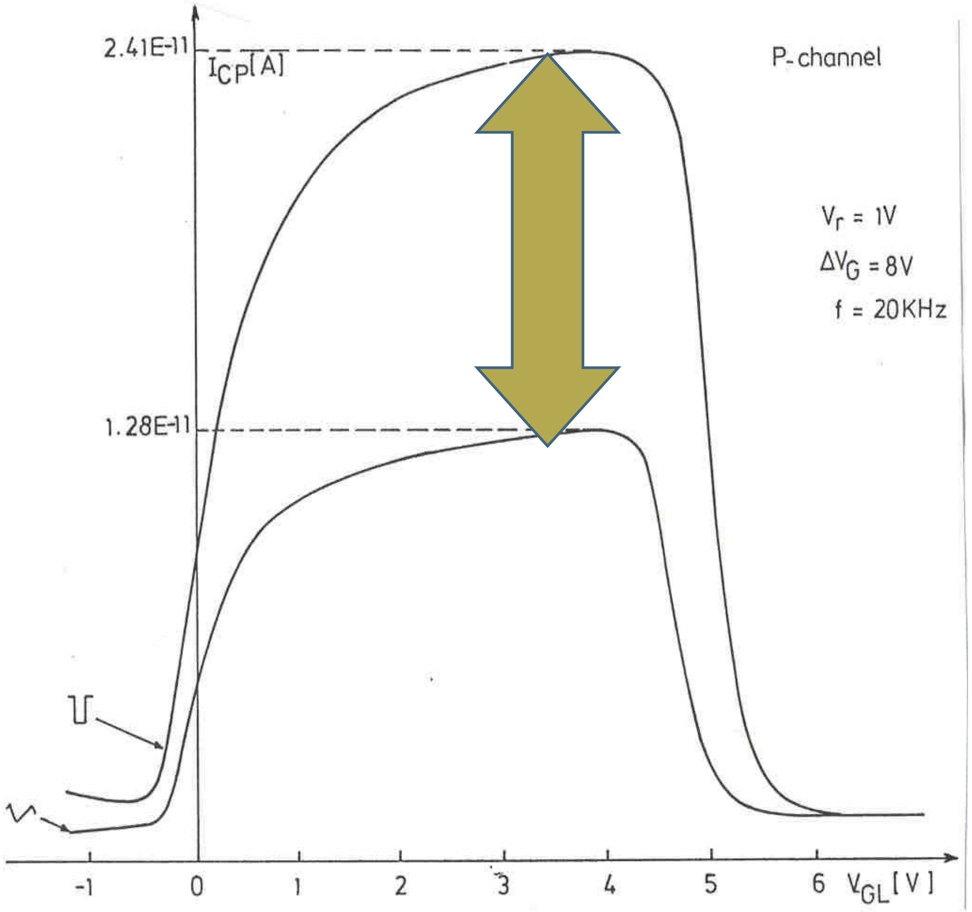
Method A

Pulse level in accumulation is fixed

Pulsing the surface into inversion with increasing amplitudes

SIMPLE OPERATING PRINCIPLE

Strong dependence on pulse shape not explained by 1st order model



Method C

Varying the pulse base level from inversion to accumulation while keeping the pulse amplitude constant

Outline

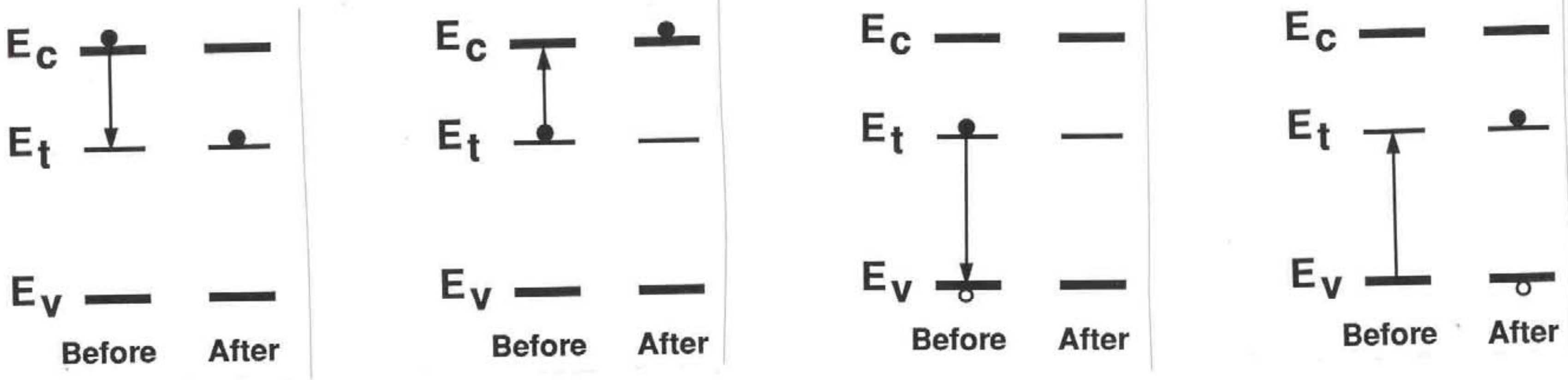
1. Introduction
2. Basic CP-principle
3. Second order model
4. Temperature dependence
5. Base level edges
6. MOSFET degradation
7. Energy distribution
8. Lateral and vertical profiling
9. Geometric components
10. Effects of oxide thickness scaling
11. Single trap characterization
12. CP in high k gate stacks

Limerick #2

Is this model correct ?

In reality there is a bit more to it
‘cause the electrons in the traps can emit
And the holes do as well
So the theory becomes a hell
And so we are in a very big sh.t

SHOCKLEY-READ-HALL THEORY



Electron capture

$N_T^0 + e^- \rightarrow N_T^-$

- capture constant
  $K_n = \sigma_n v_{th}$ ($cm^3 s^{-1}$)
- capture rate
  $c_n = K_n n$ (s^{-1})

Electron emission

$N_T^- \rightarrow N_T^0 + e^-$

- emission rate e_n (s^{-1})

Hole capture

$N_T^- + h^+ \rightarrow N_T^0$

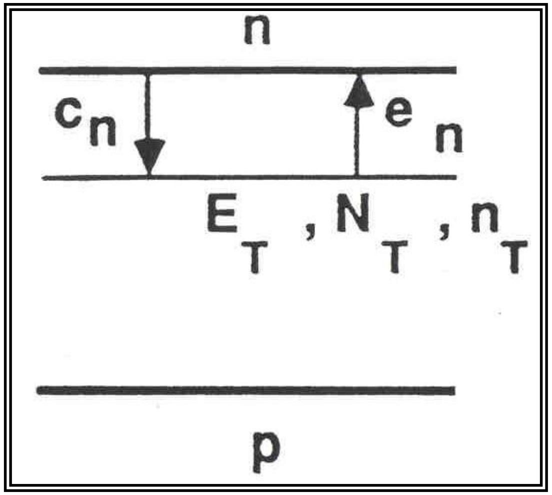
- capture constant
  $K_p = \sigma_p v_{th}$ ($cm^3 s^{-1}$)
- capture rate
  $c_p = K_p p$ (s^{-1})

Hole emission

$N_T^0 \rightarrow N_T^- + h^+$

- emission rate e_p (s^{-1})

SHOCKLEY-READ-HALL THEORY



Detailed balance in conduction band:

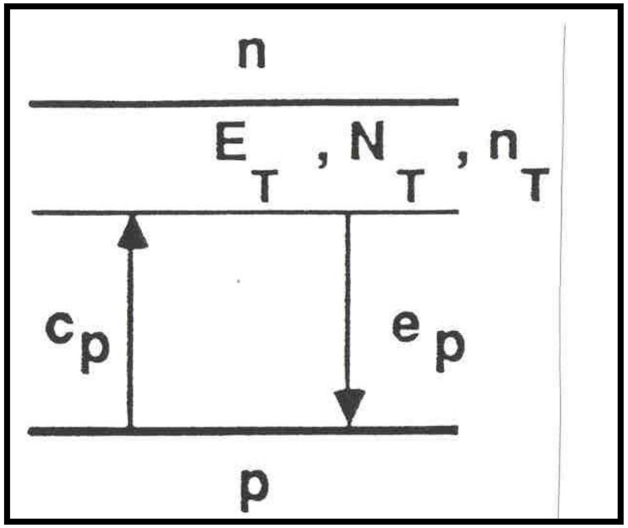
$$\frac{dn}{dt} = e_n n_T - K_n n(N_T - n_T)$$

with $n_T = N_T f_T$ and $f_T = \left[1 + \exp\frac{(E_T - E_F)}{kT}\right]^{-1}$

In equilibrium: $\frac{dn}{dt} = 0 \implies e_n = \frac{K_n n(N_T - n_T)}{n_T} = K_n n \left(\frac{1}{f_T} - 1\right)$ with $n = n_i \exp\left(\frac{E_F - E_i}{kT}\right)$

$$e_n = \sigma_n v_{th} n_i \exp\left(\frac{E_T - E_i}{kT}\right)$$

SHOCKLEY-READ-HALL THEORY



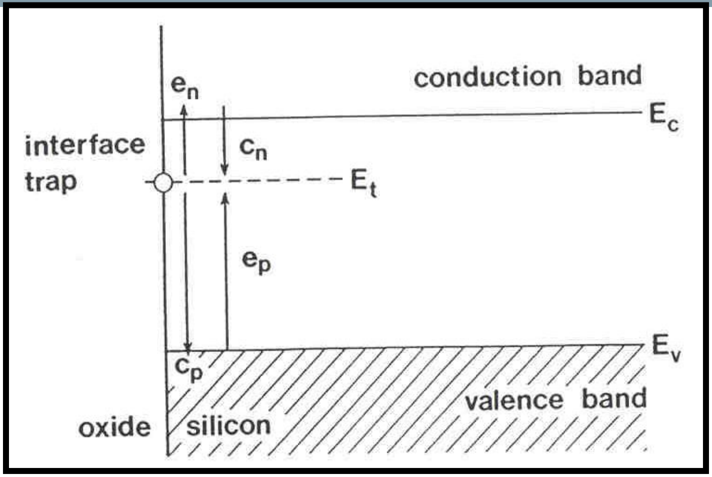
Detailed balance in valence band:

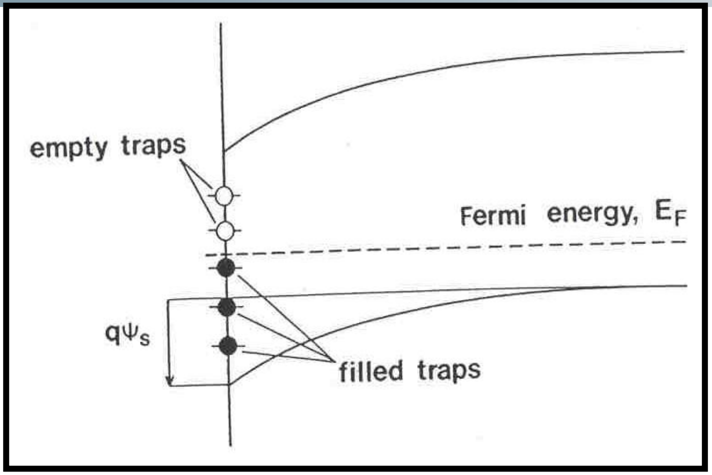
$$\frac{dp}{dt} = e_p \left(N_T - n_T \right) - K_p p \, n_T$$

$$\text{In equilibrium}: \frac{dp}{dt} = 0 \implies e_p = \frac{K_p p n_T}{(N_T - n_T)} = K_p \, p \left(\frac{1}{f_T} - 1 \right)^{-1} \quad \text{with} \quad p = n_i \exp\left(\frac{E_i - E_F}{kT} \right)$$

$$e_p = \sigma_p \, v_{th} \, n_i \exp\left(\frac{E_i - E_T}{kT} \right)$$

SHOCKLEY-READ-HALL THEORY

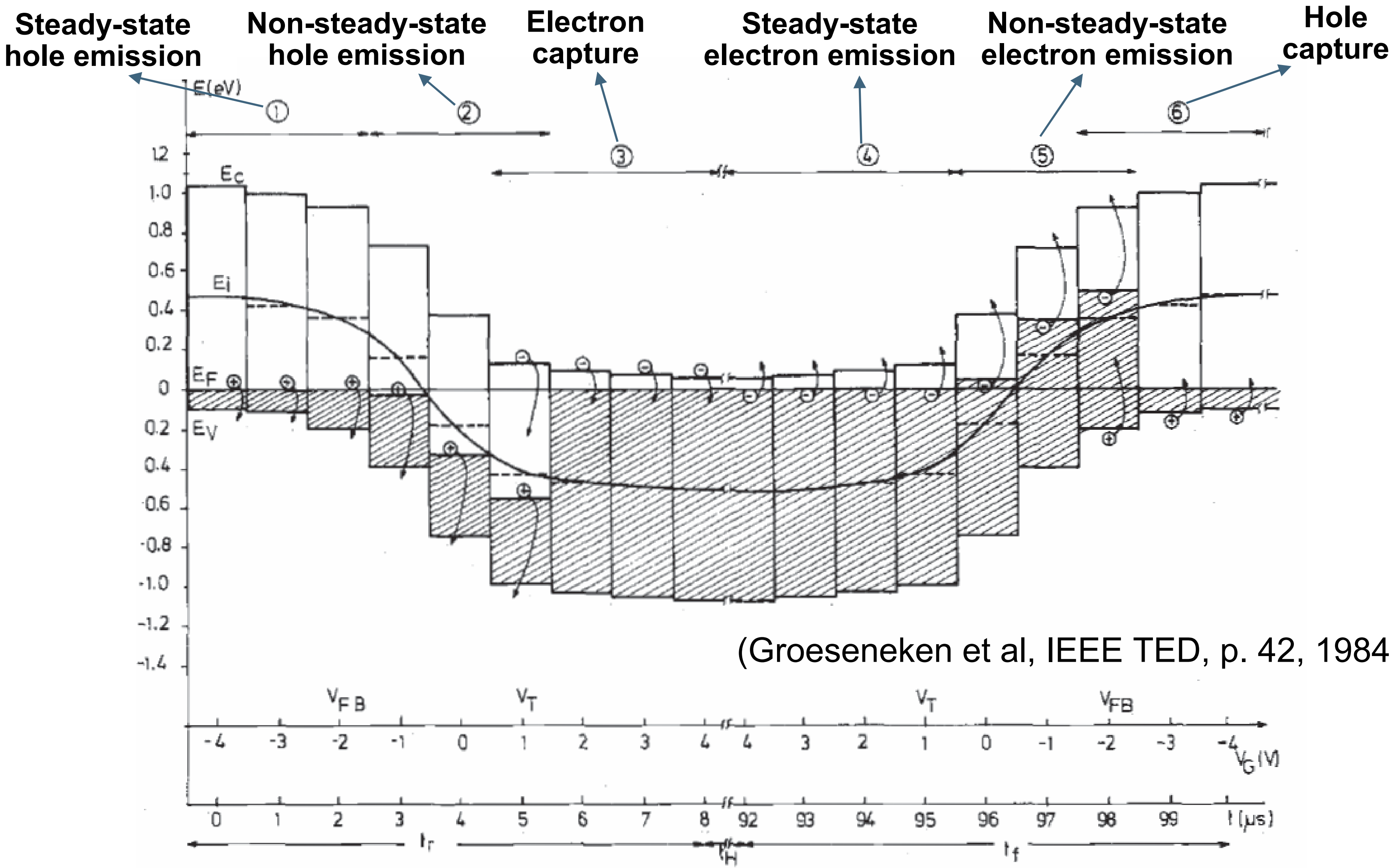




$$\frac{df_T(E_T,t)}{dt} = (c_n + e_p(E_T)\ (1 - f_T(E_T,t) - (e_n(E_T) + c_p) f_T(E_T,t)$$

- Electron emission to conduction band: $e_n = \sigma_n v_{th} n_i \exp\left(\frac{E_T - E_i}{kT}\right)$
- Electron capture from conduction band: $c_n = \sigma_n v_{th} n = \sigma_n v_{th} n_i \exp\left(\frac{E_F - E_i}{kT}\right)$
- Hole emission to valence band: $e_p = \sigma_p v_{th} n_i \exp\left(\frac{E_i - E_T}{kT}\right)$
- Hole capture from valence band: $c_p = \sigma_p v_{th} p = \sigma_p v_{th} n_i \exp\left(\frac{E_i - E_F}{kT}\right)$

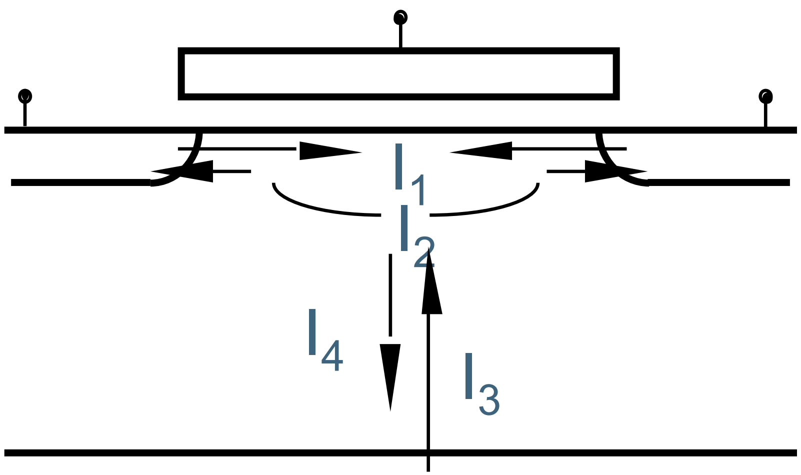
Charge pumping : 2nd order understanding



(Groeseneken et al, IEEE TED, p. 42, 1984)

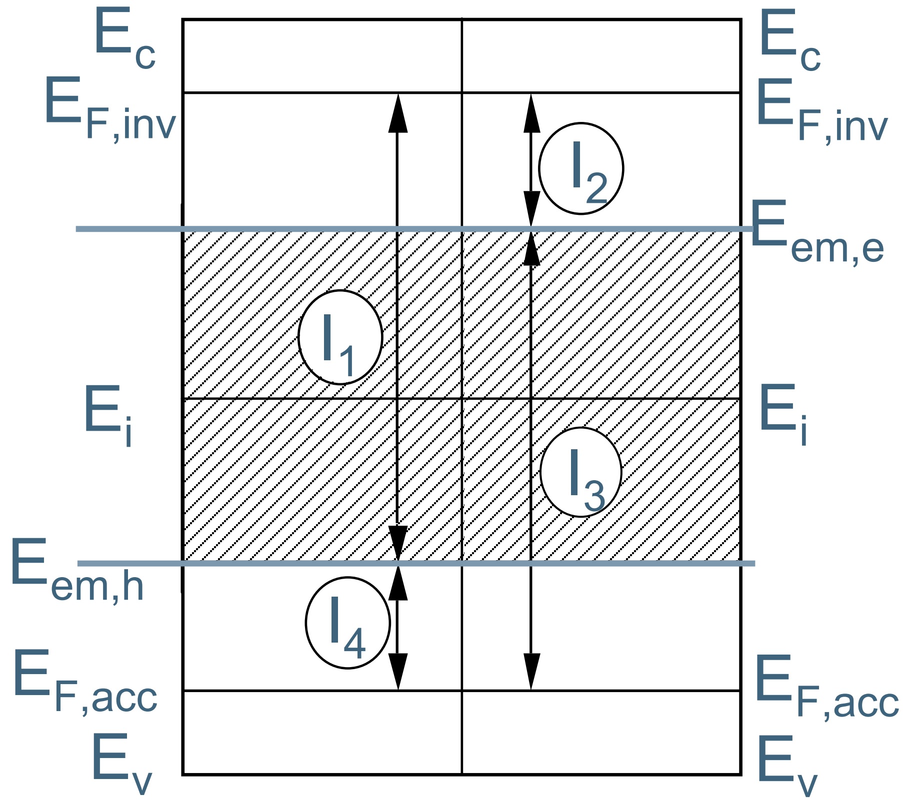
Interface trap processes occurring during one CP-cycle

Thermal emission/recombination



I_1 = trapping of electrons
I_2 = emission of electrons
I_3 = trapping of holes
I_4 = emission of holes

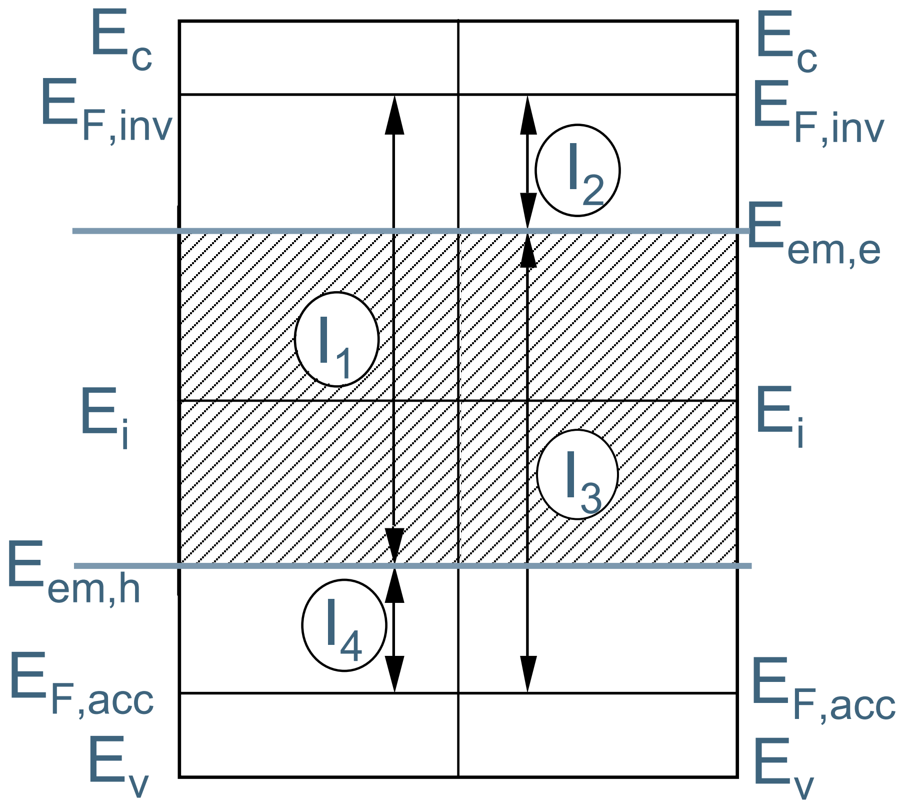
$$I_{CP} = I_3 - I_4 = I_1 - I_2$$
$$= q\,f\,A_G\,D_{it}\,(E_{em,e} - E_{em,h})$$



Different energy regions associated with the four current components

(Groeseneken et al, IEEE TED, p. 42, 1984)

Thermal emission/recombination



$$I_{cp} = q\, f\, A\, D_{it}\, (E_{em,e} - E_{em,h})$$

From SRH emission theory it can be shown that Eem,h and Eem,e are given by:

$$E_{em,h}(t) = E_i + kT\ \ln(\sigma_p\, v_{th}\, n_i\, t_{em,h})$$

$$E_{em,e}(t) = E_i - kT\ \ln(\sigma_n\, v_{th}\, n_i\, t_{em,e})$$

<u>Assumptions</u>:

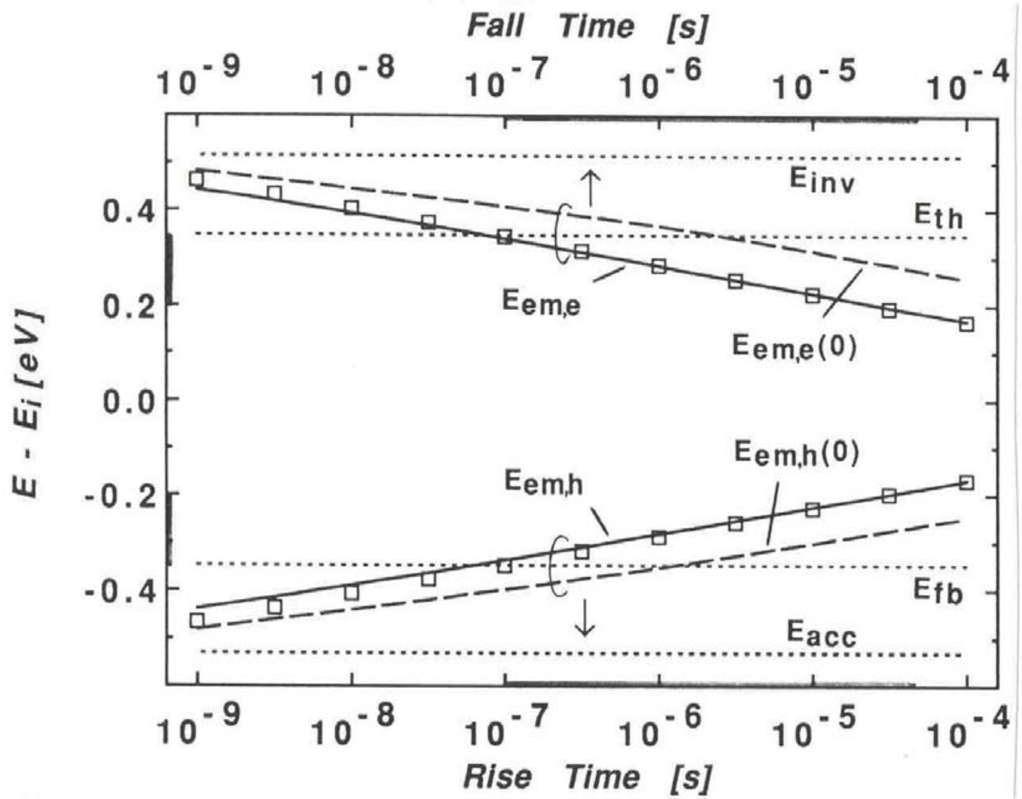
- n.s.s. emission times are sufficiently long
- n.s.s. emission occurs only if $V_{FB}<V_G<V_T$

$$t_{em,h} = \frac{|V_{FB} - V_T|}{\Delta V_A} \times t_r$$

$$t_{em,e} = \frac{|V_{FB} - V_T|}{\Delta V_A} \times t_f$$

(Groeseneken et al, IEEE TED, p. 42, 1984)

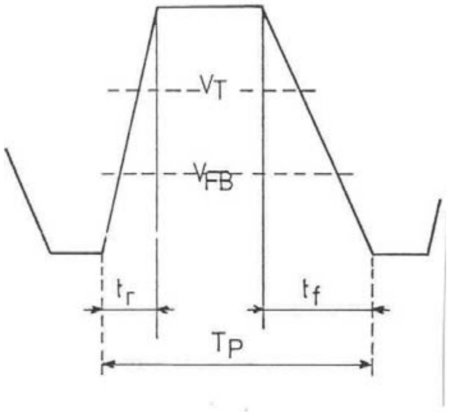
CHARGE PUMPING THEORY



Dashed lines:	$E_{em,h}$ (0) and $E_{em,e}$(0)
Solid lines:	Exact values of $E_{em,h}$ and $E_{em,e}$
Squares:	Approximations for $E_{em,h}$ and $E_{em,e}$

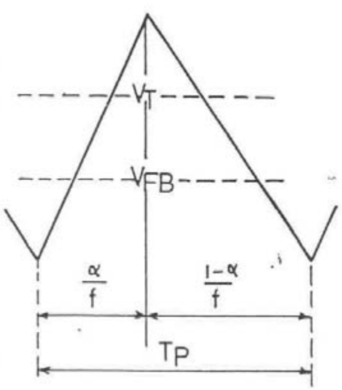
Pulse transient time dependence of the emission levels

Charge pumping expressions



Square pulses :

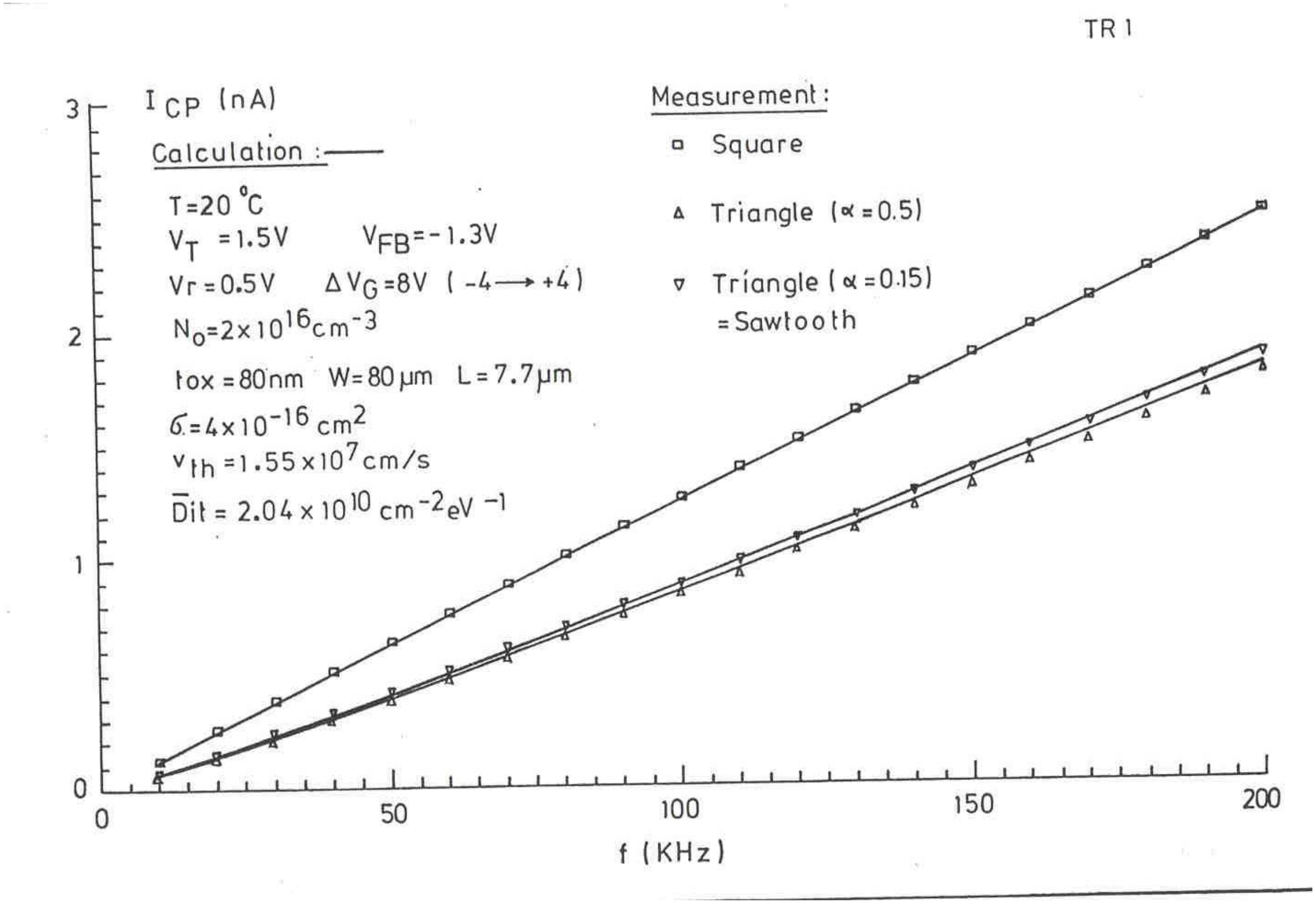
$$I_{cp} = 2 \cdot q \cdot D_{it} \cdot f \cdot A_G \cdot kT \cdot \ln\left(v_{th} n_i \sqrt{\sigma_n \sigma_p} \frac{|V_{fb} - V_t|}{\Delta V_G} \sqrt{t_r t_f} \right)$$



Triangular pulses :

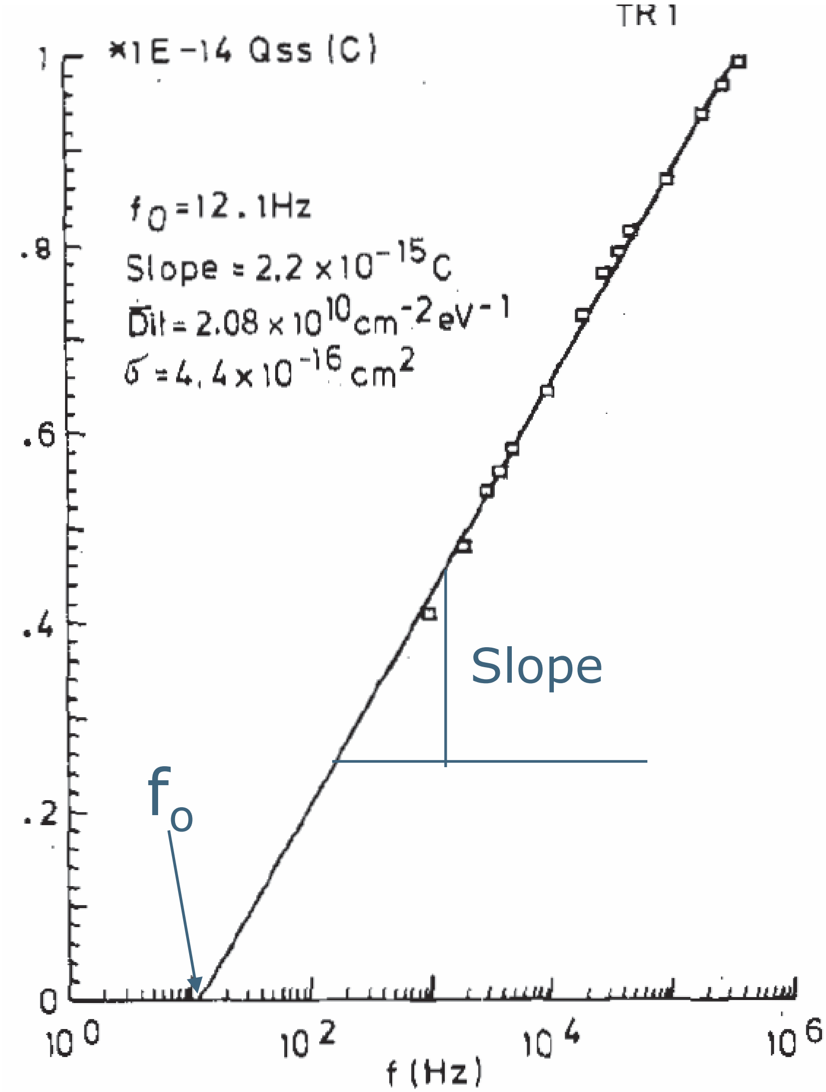
$$I_{cp} = 2 \cdot q \cdot D_{it} \cdot f \cdot A_G \cdot kT \cdot \ln\left(v_{th} n_i \sqrt{\sigma_n \sigma_p} \frac{|V_{fb} - V_t|}{\Delta V_G} \frac{\sqrt{\alpha (1 - \alpha)}}{f} \right)$$

Frequency dependence



Frequency dependence of I_{CP} for square pulses and triangular pulses with α = 0.5 and α = 0.15.

Thermal emission/recombination



For triangular pulses:

$Q_{cp} = I_{cp}/f$ vs. log(f) is a straight line

$$\sqrt{\sigma_n \sigma_p} = \frac{1}{v_{th} n_i} \cdot \frac{\Delta V_A}{|V_t - V_{fb}|} \cdot 2f_o$$

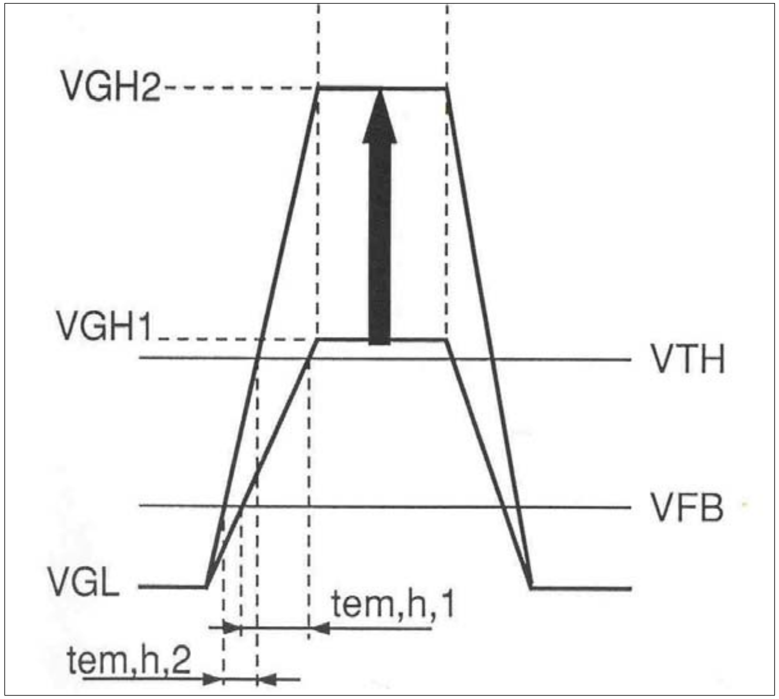
$$D_{it} = \frac{\log e}{2 \cdot q \cdot kT \cdot A_G} \cdot \text{Slope}$$

(Groeseneken et al, IEEE TED, p. 42, 1984)

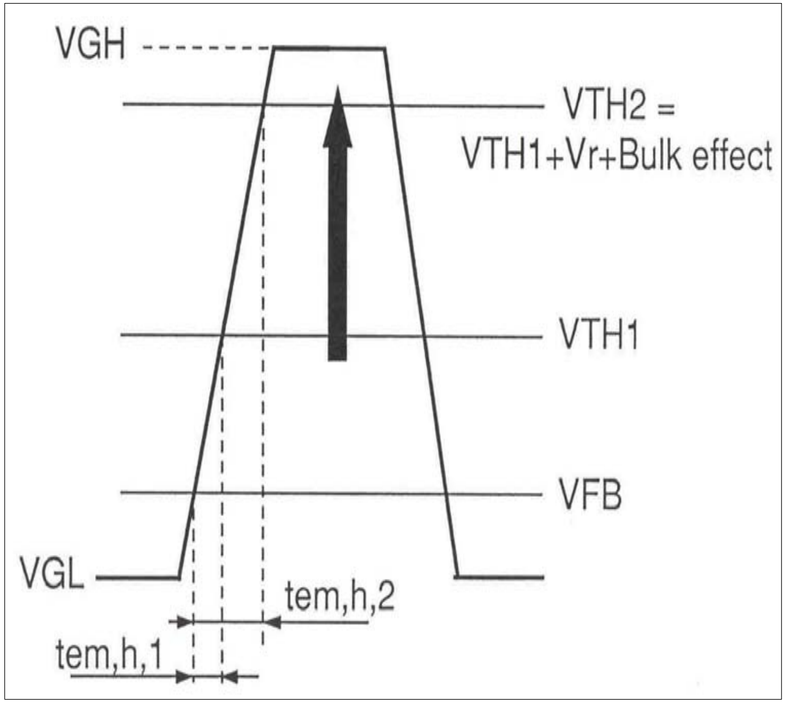
EXPERIMENTAL RESULTS

Influence of

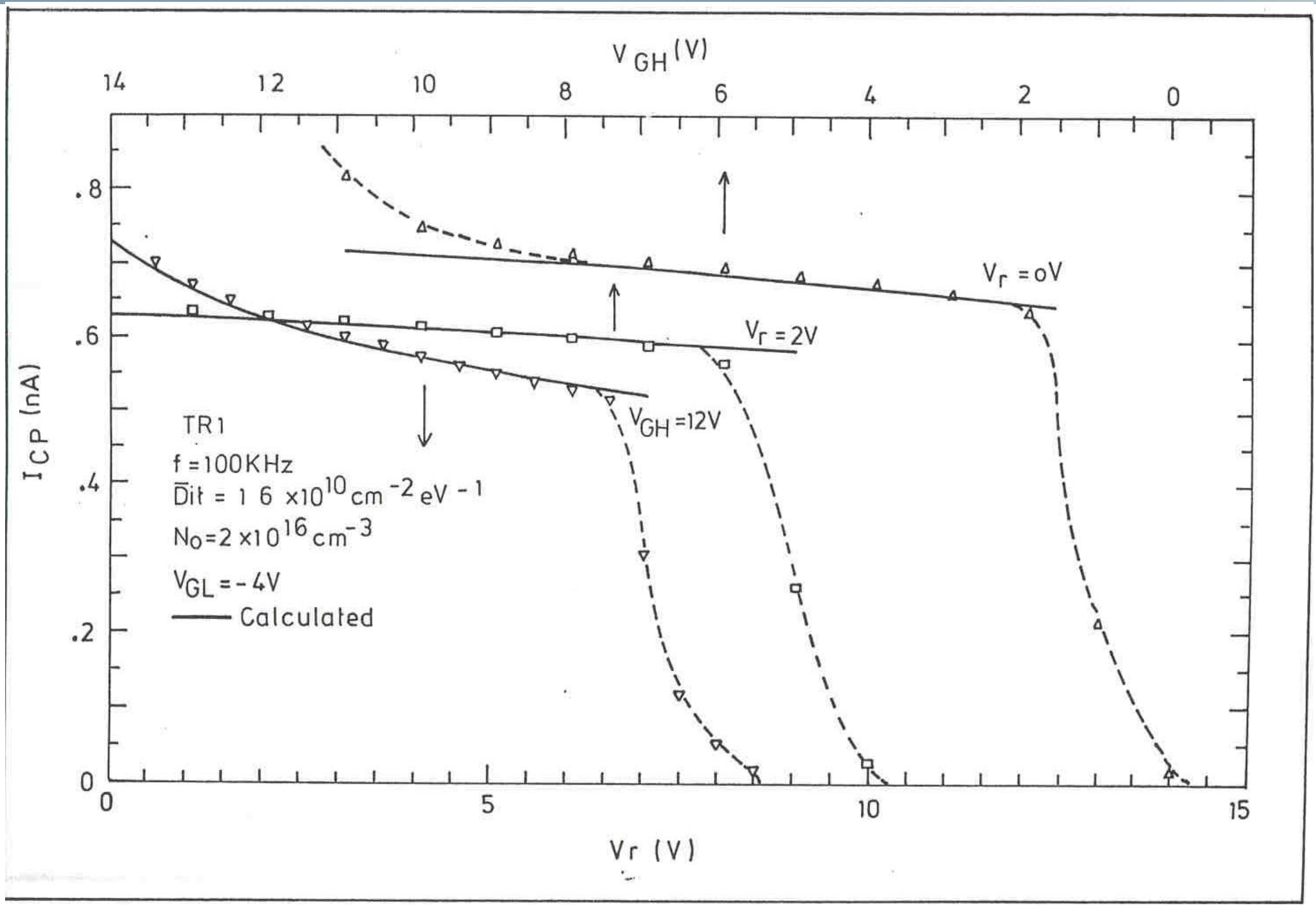
High level of gate voltage pulse



Reverse voltage at source & drain



Amplitude and reverse voltage dependence

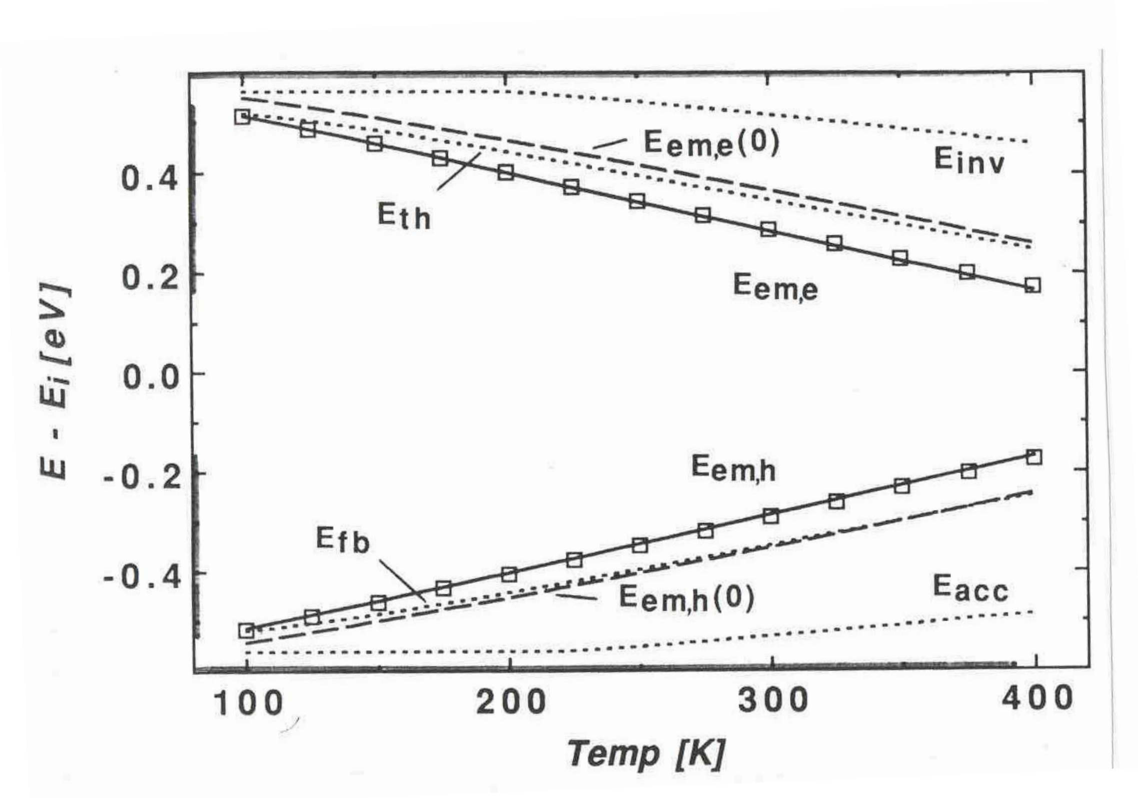


Influence of high level of gate voltage pulse and of reverse voltage at source and drain

Outline

1. Introduction
2. Basic CP-principle
3. Second order model
4. Temperature dependence
5. Base level edges
6. MOSFET degradation
7. Energy distribution
8. Lateral and vertical profiling
9. Geometric components
10. Effects of oxide thickness scaling
11. Single trap characterization
12. CP in high k gate stacks

CHARGE PUMPING THEORY



Dashed lines: $E_{em,h}$ (0) and $E_{em,e}$(0)

Solid lines: Exact values of $E_{em,h}$ and $E_{em,e}$

Squares: Approximations for $E_{em,h}$ and $E_{em,e}$

Temperature dependence of the emission levels

TEMPERATURE DEPENDENCE

- Temperature depence of I_{CP} is of the general form:

$$I_{CP}(T) = -aT - bTLnT + c$$

with

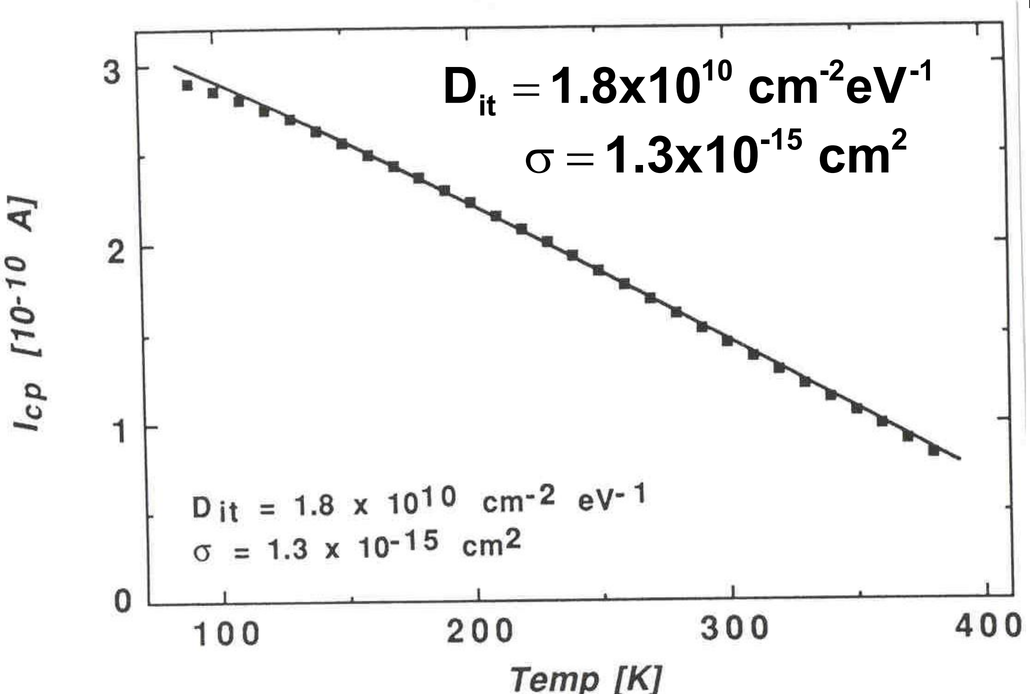
$$a = 2qkfA_G D_{it} \ln\left(\sigma\sqrt{\frac{3k}{m^*}} K_i \sqrt{t_{em,e} t_{em,h}}\right)$$

$$b = 4qkfA_G D_{it}$$

$$c = qfA_G D_{it} E_g$$

$$D_{it} = 1.8\text{x}10^{10}\ \text{cm}^{-2}\text{eV}^{-1}$$

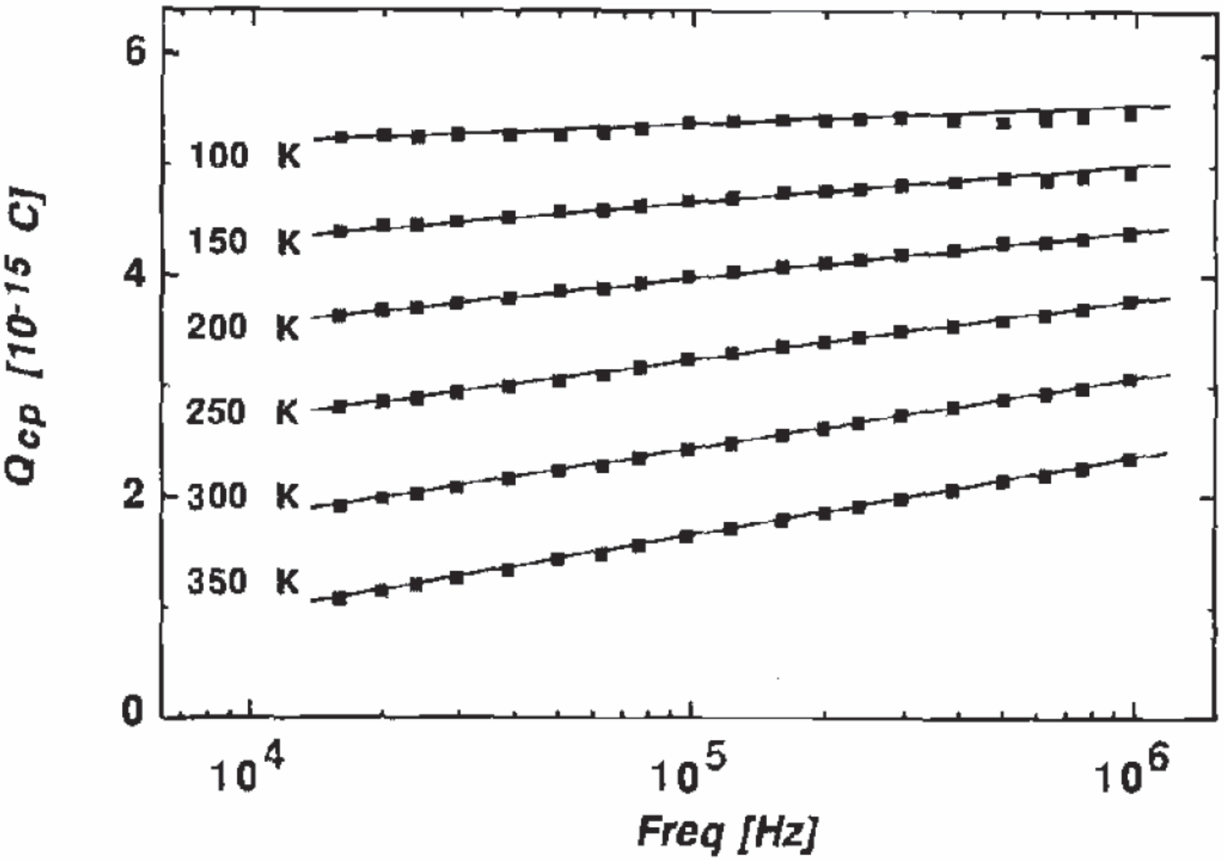
$$\sigma = 1.3\text{x}10^{-15}\ \text{cm}^2$$



Measured data can be fitted to this analytical expression yielding a constant D_{it} and σ in the temperature range from 90K to 380K

(Van den bosch et al., IEEE TED, p. 1820, 1991)

Dependence on temperature



$$f_o = \frac{\sigma v_{th}(T) n_i(T) \left| V_{TH} - V_{FB} \right|}{2\Delta V_G}$$

$$\ln(f_o) = A + 2\ln T - \frac{E_g}{2kT}$$

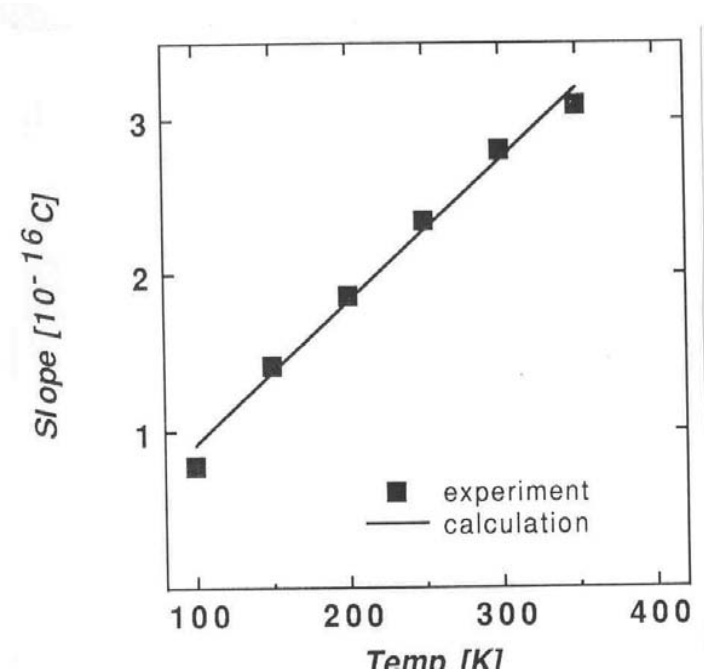
$$\text{with } A = \ln\left(\sigma \sqrt{\frac{3k}{m^*}} K_i \frac{\left| V_{TH} - V_{FB} \right|}{2\Delta V_G} \right)$$

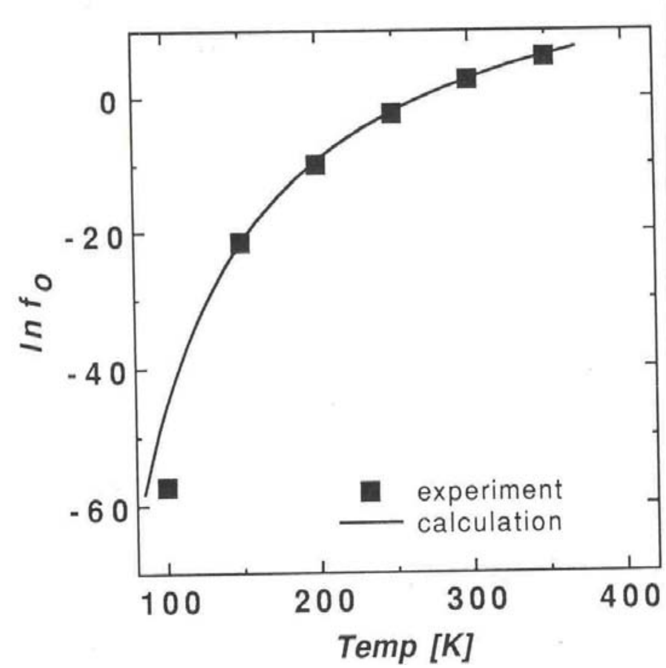
$$\text{Slope} = \frac{dQ_{CP}}{d\ln f} = 2\,q\,kT\,A_G\,D_{it}$$

(Van den bosch et al., IEEE TED, p. 1820, 1991)

At low T, less thermal emission, hence more recombination and higher I_{cp}

TEMPERATURE DEPENDENCE





$$\text{Slope} = \frac{dQ_{CP}}{d\ln f} = 2\,q\,kT\,A_G\,D_{it}$$

$$\ln(f_o) = \ln\left(\sigma\sqrt{\frac{3k}{m^*}}K_i\,\frac{|V_{TH} - V_{FB}|}{2\Delta\Delta_G}\right) + 2\ln T - \frac{Eg}{2kT}$$

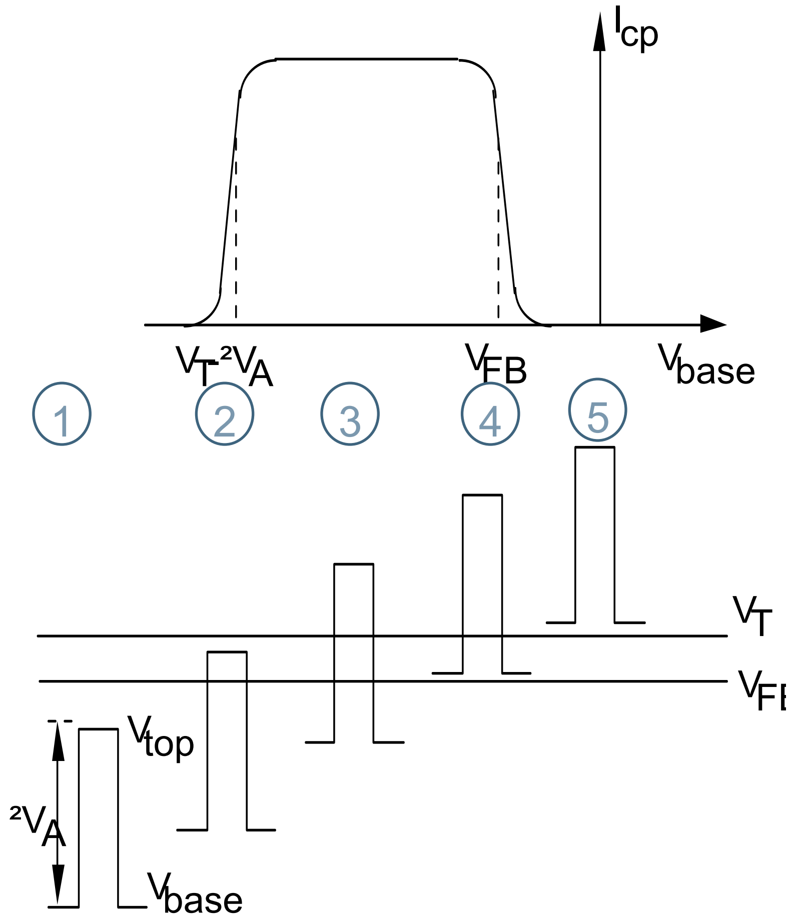
$$D_{it} = 1.7 \times 10^{10}\,cm^{-2}eV^{-1}$$

$$\sigma = 1\times10^{15}cm^2$$

Outline

1. Introduction
2. Basic CP-principle
3. Second order model
4. Temperature dependence
5. Base level edges
6. MOSFET degradation
7. Energy distribution
8. Lateral and vertical profiling
9. Geometric components
10. Effects of oxide thickness scaling
11. Single trap characterization
12. CP in high k gate stacks

Method C: Base level technique



Region 1: Vbase<Vtop<Vfb<Vt
no CP-current

Region 2: Vbase<Vfb<Vtop<Vt
transition from 0 to Icp,max

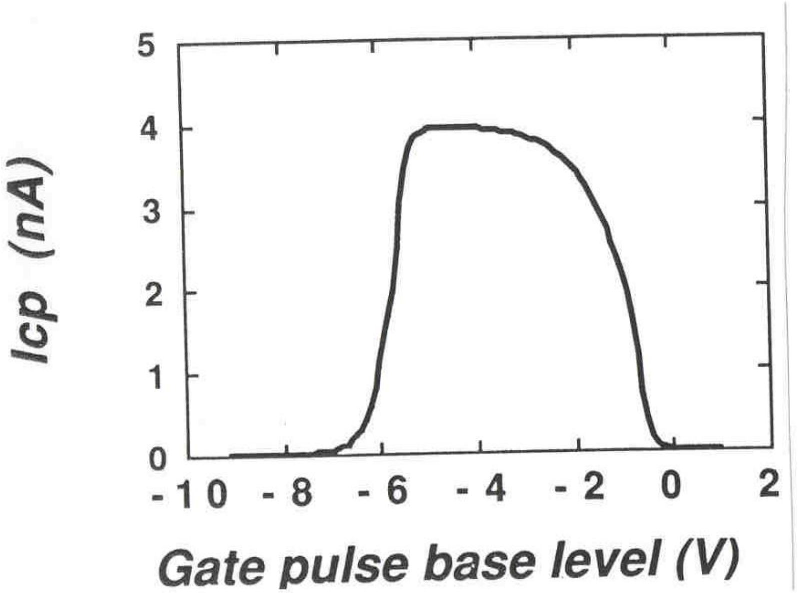
Region 3: Vbase<Vfb<Vt<Vtop
normal CP-regime

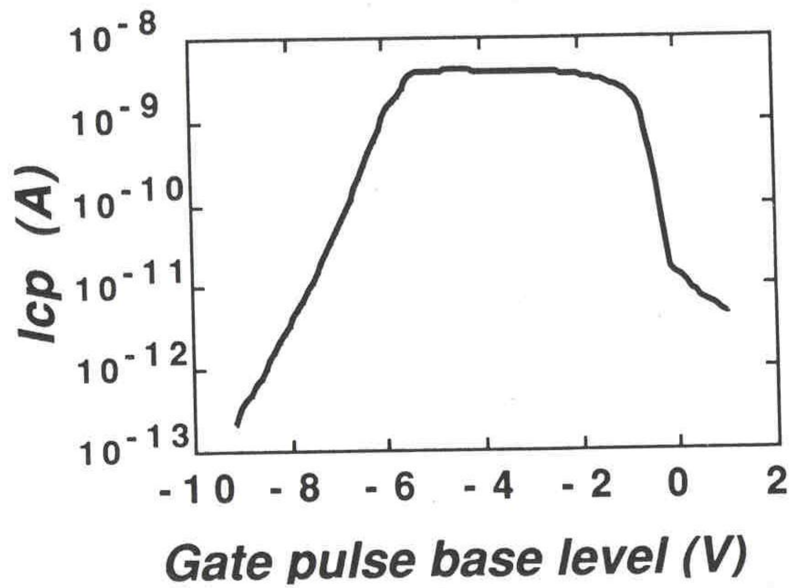
Region 4: Vfb<Vbase<Vt<Vtop
transition form Icp,max to 0

Region 5: Vfb<Vt<Vbase<Vtop

Principle of the base level technique

IMPROVEMENTS TO THE MODEL





W/L = 100μm / 2μm
Freq = 1MHz
Vr = 0.1V

- Experimental n-channel charge pumping characteristics

IMPROVEMENTS TO THE MODEL

New definition of threshold and flatband voltage for charge pumping measurements

$$\tau_{cn} = \frac{1}{\sigma_n v_{th} n_s} \qquad \tau_{cp} = \frac{1}{\sigma_p v_{th} p_s}$$

$\mathbf{V_t}$ **=** V_g where n_s is large enough for electrons to be captured in fast interface traps during the high part of the gate pulse: V_m (minority carriers)

$\mathbf{V_{fb}}$ **=** V_g where p_s is large enough for holes to be captured in fast interface traps during the low part of the gate pulse: V_M (majority carriers)

$$\mathbf{n_s = \frac{4f}{\sigma_n v_{th}}} \qquad \mathbf{p_s = \frac{4f}{\sigma_p v_{th}}}$$

$$\text{e.g. } f = 100\text{kHz}, \sigma_n = 2\times10^{-15}\text{cm}^2 \rightarrow n_s = 2\times10^{14}\text{cm}^{-3}$$

IMPROVEMENTS TO THE MODEL

Conventional definition:

$$V_t = V_g\left(\Psi_s = 2\varphi_F = 2\frac{kT}{q}\ln\left(\frac{N}{n_i}\right)\right)$$

$$V_{fb} = V_g(\Psi_s = 0)$$

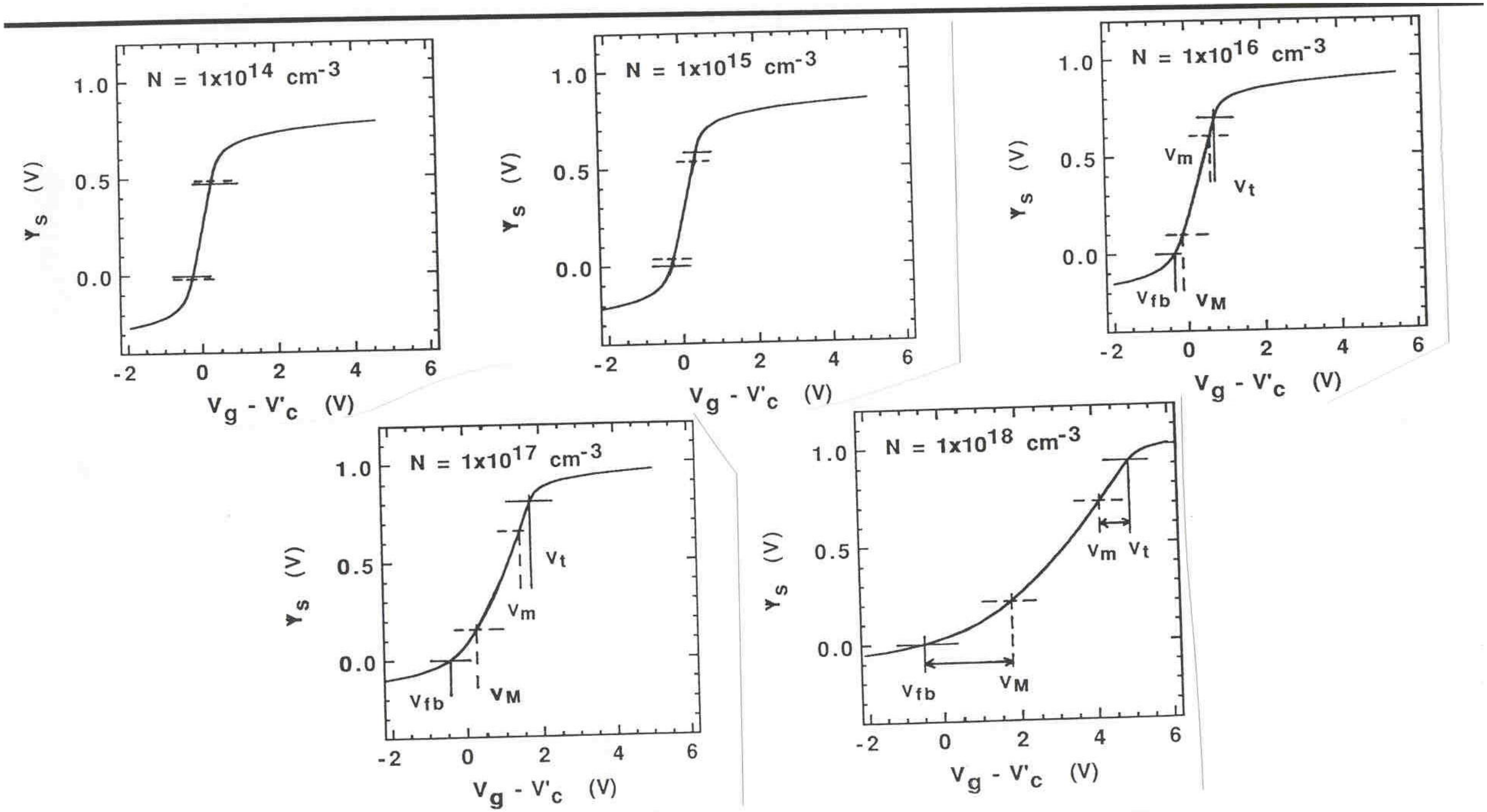
CP-definition:

$$n_s = \frac{n_i^2}{N}\exp(q\,\Psi_s/kT) = \frac{4f}{v_{th}\sigma_n} \rightarrow V_m = V_g\left(\Psi_s = \frac{kT}{q}\ln\left(\frac{N}{n_i^2}\frac{4f}{v_{th}\sigma_n}\right)\right)$$

$$p = N\exp(-q\,\Psi_s/kT) = \frac{4f}{v_{th}\sigma_p} \rightarrow V_M = V_g(\Psi_s = \frac{kT}{q}\ln\left(\frac{Nv_{th}\sigma_p}{4f}\right))$$

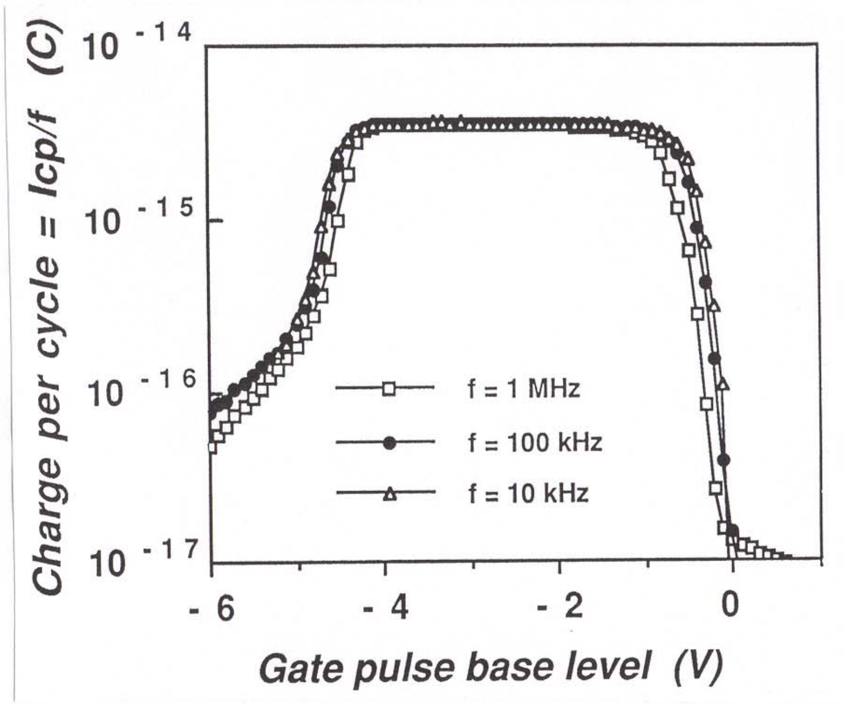
Comparison with conventional definitions of V_t and V_{fb}

IMPROVEMENTS TO THE MODEL

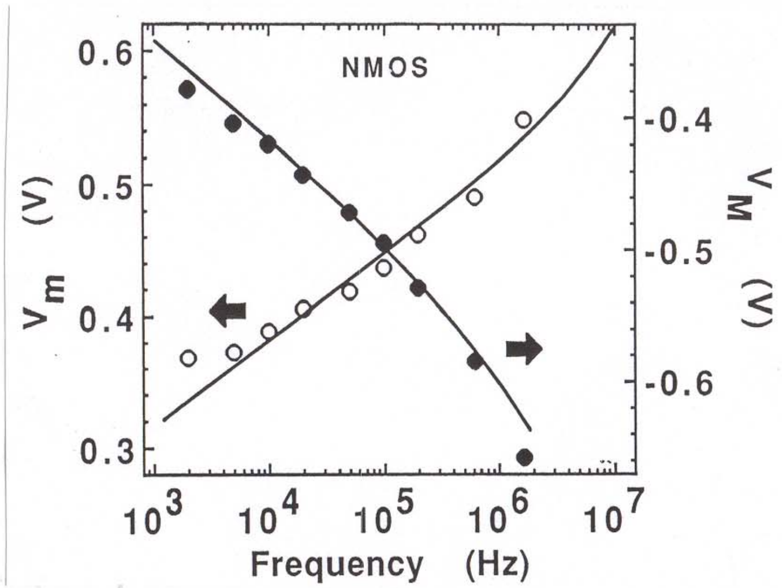


Difference between V_t and V_m and between V_{fb} and V_M for various doping levels N

IMPROVEMENTS TO THE MODEL

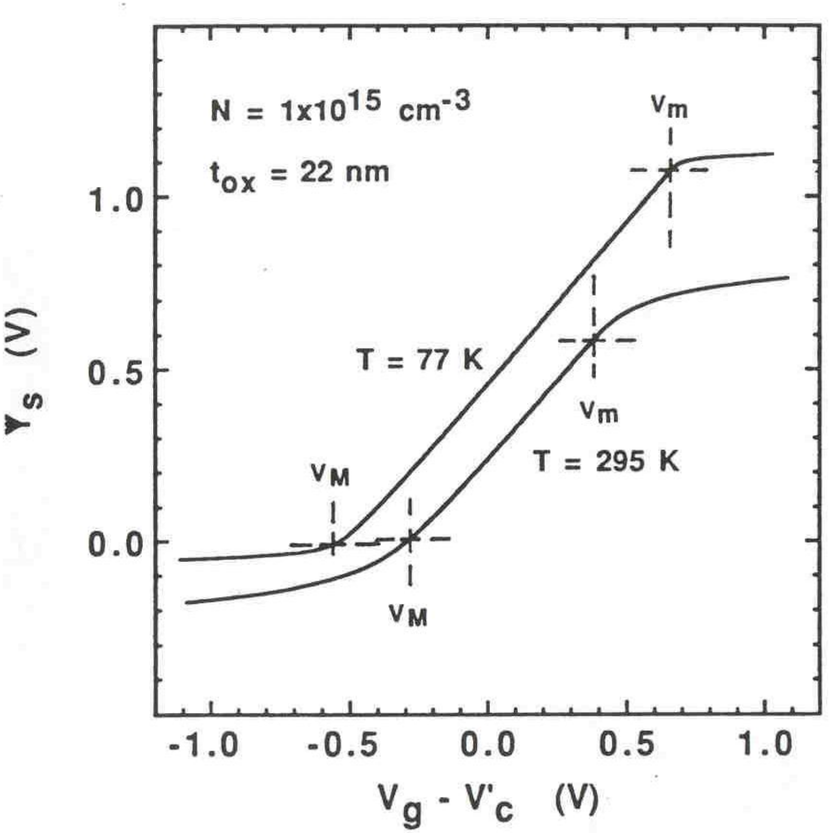


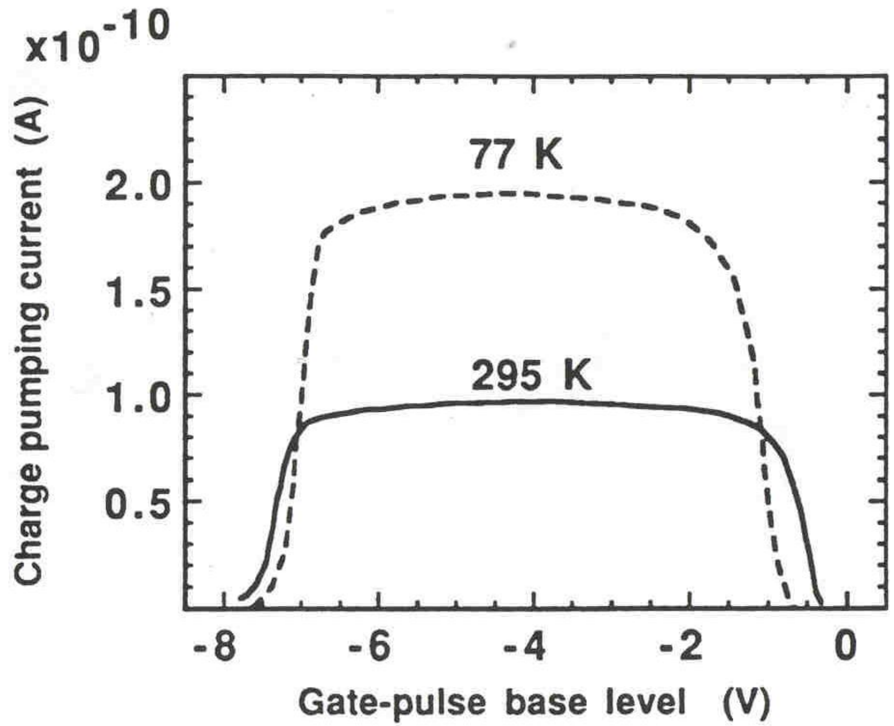
Charge per cycle vs base level at various pulse frequencies



Vm and VM vs frequency
Symbols = experiment
Solid lines = theory

Frequency dependence of “threshold” (V_m) and “flatband” (V_M) voltage

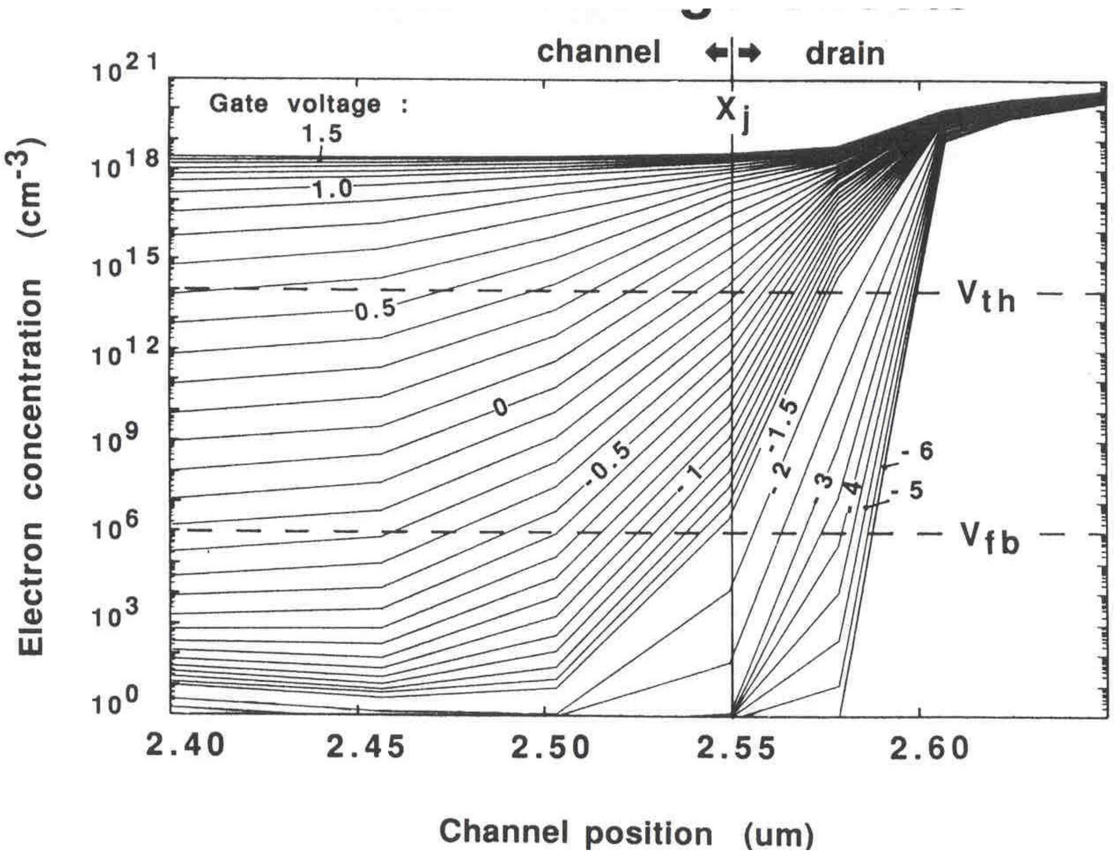




Influence of temperature on V_m and V_M

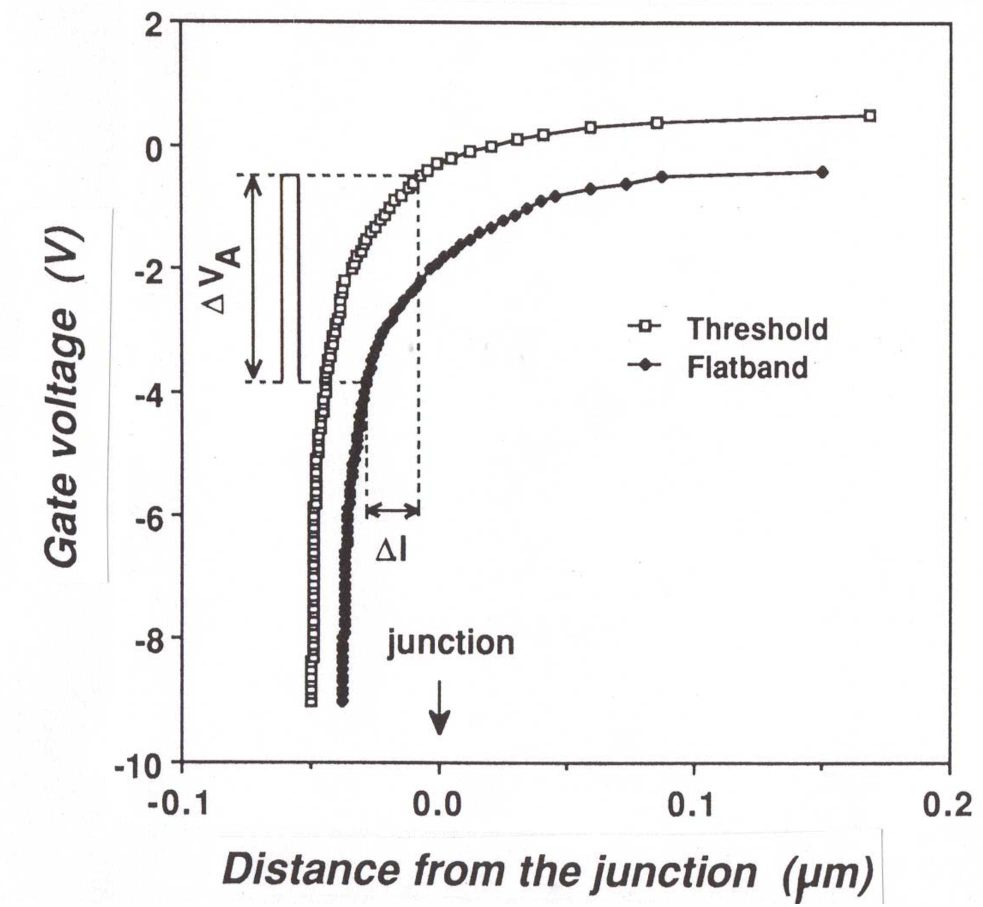
IMPROVEMENTS TO THE MODEL

Influence of edge effects



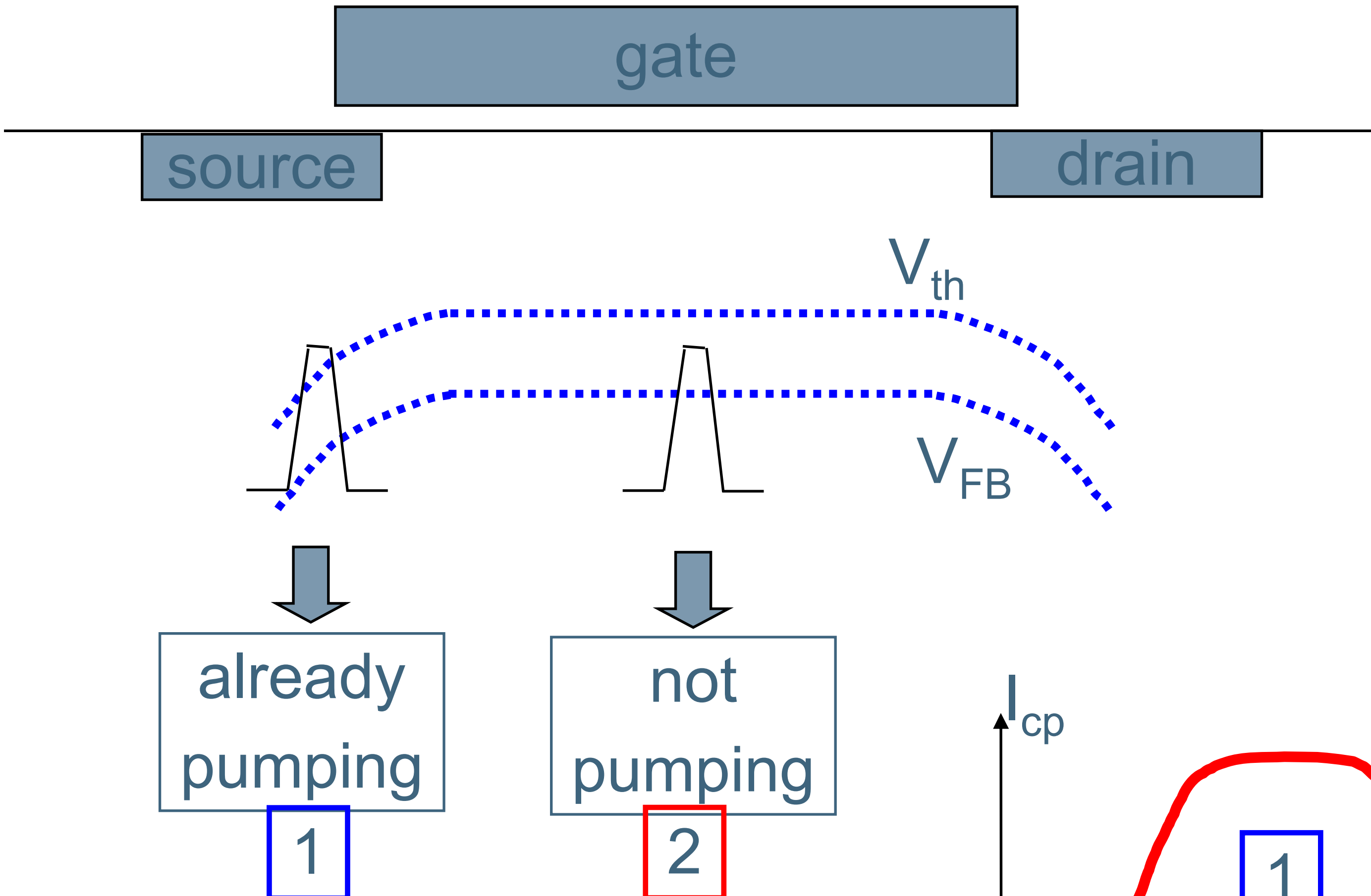
- Determination of V_t and V_{fb} using MINIMOS

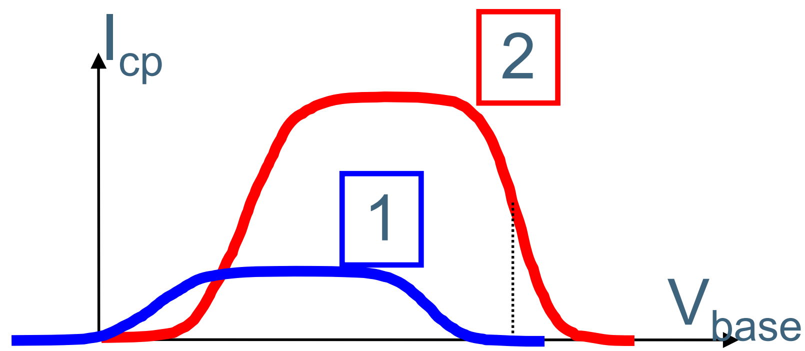
IMPROVEMENTS TO THE MODEL



- Spatial dependence of V_t and V_{fb} near source and drain

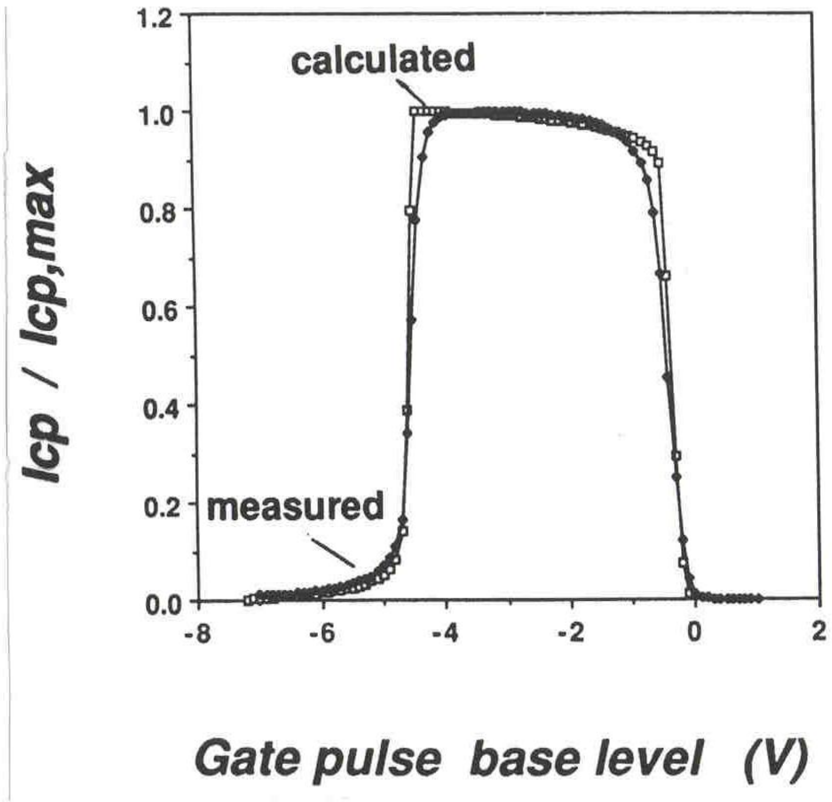
Influence of device edges

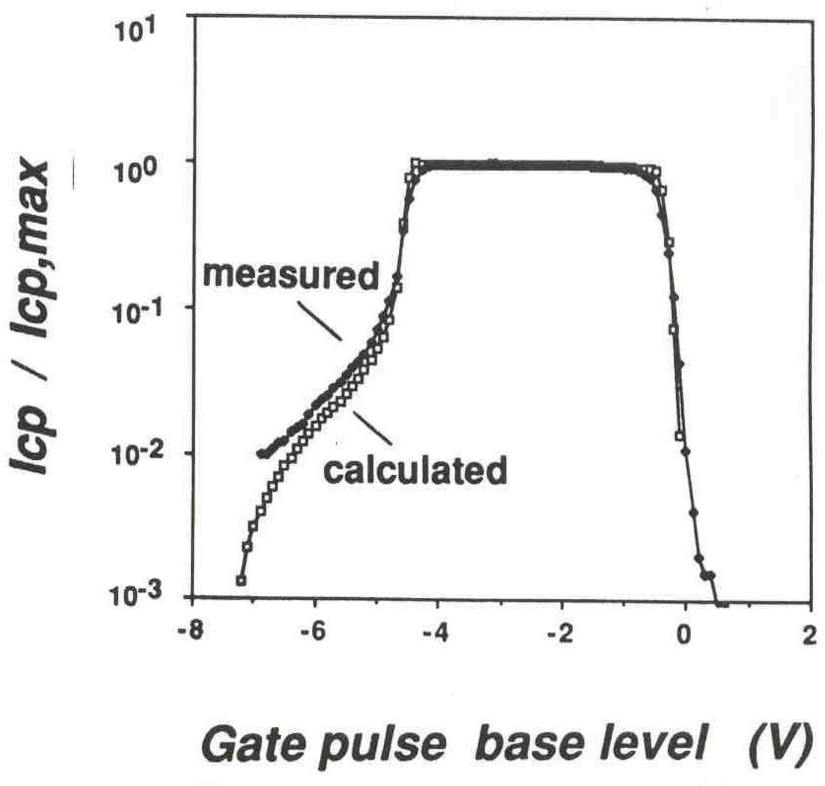




(Heremans et al., IEEE TED, p. 1318, 1989)

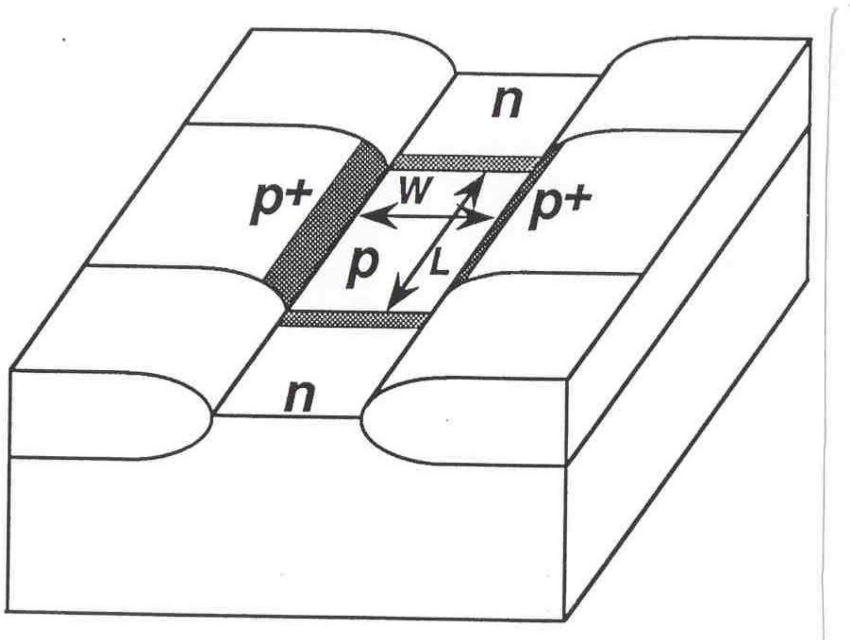
IMPROVEMENTS TO THE MODEL

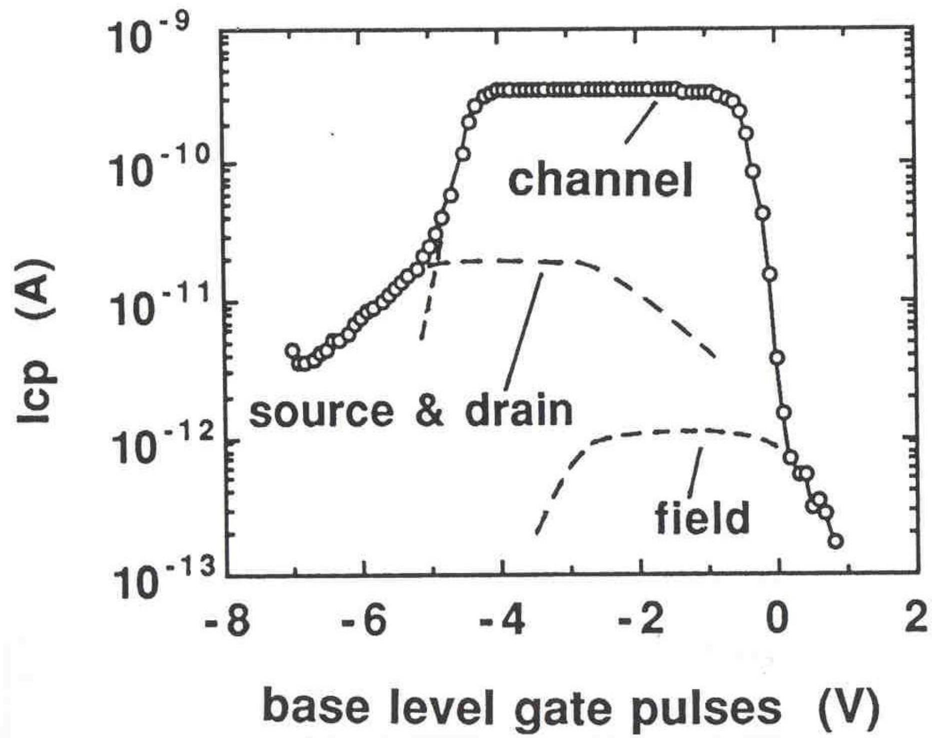




Influence of edge effects at source and drain:
theory versus experiment

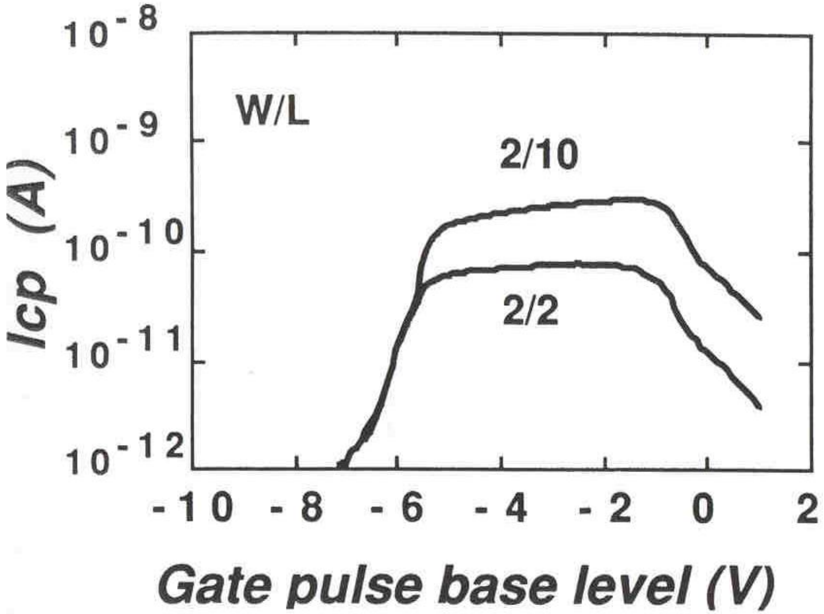
Influence of device edges

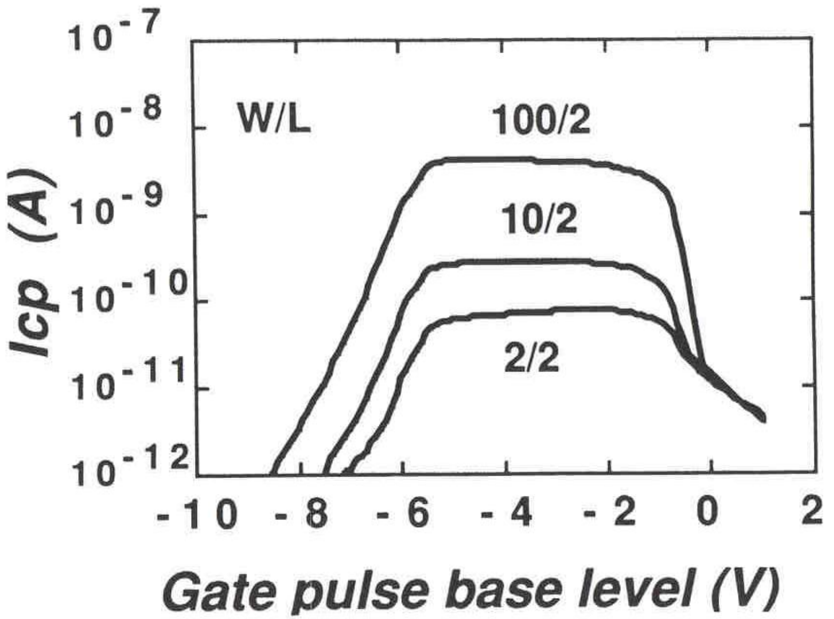




Edges of base level curve contain information on the interface characteristics at S and D and at the field edges

Influence of device edges





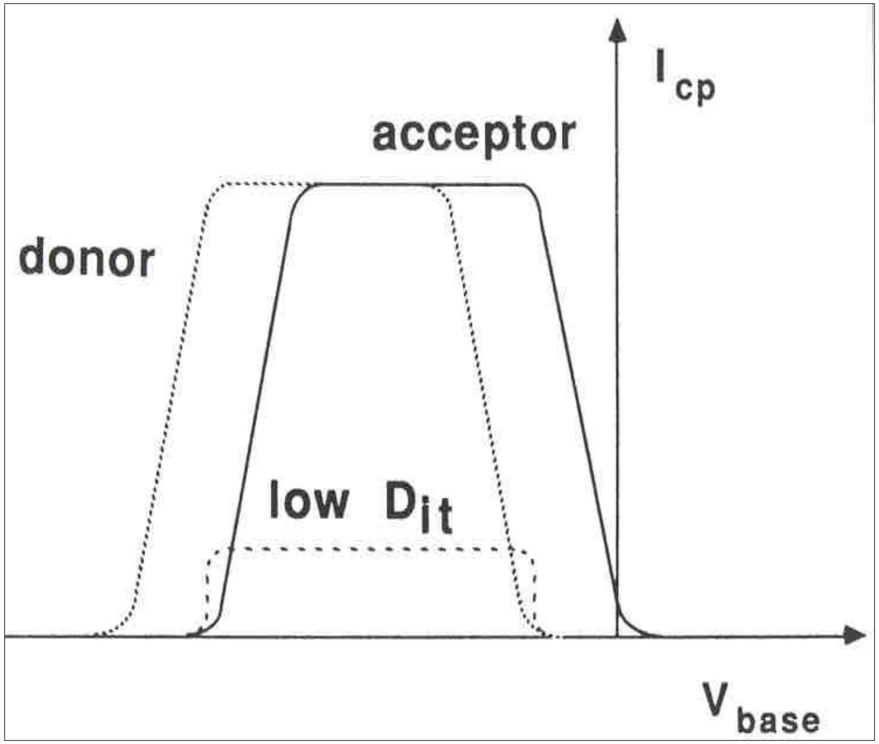
Fixed channel width
Varying channel length

LOCOS edge sensitive

Fixed channel length
Varying channel width

S/D edge sensitive

Influence of interface trapped charge



for t_{ox} = 20 nm:

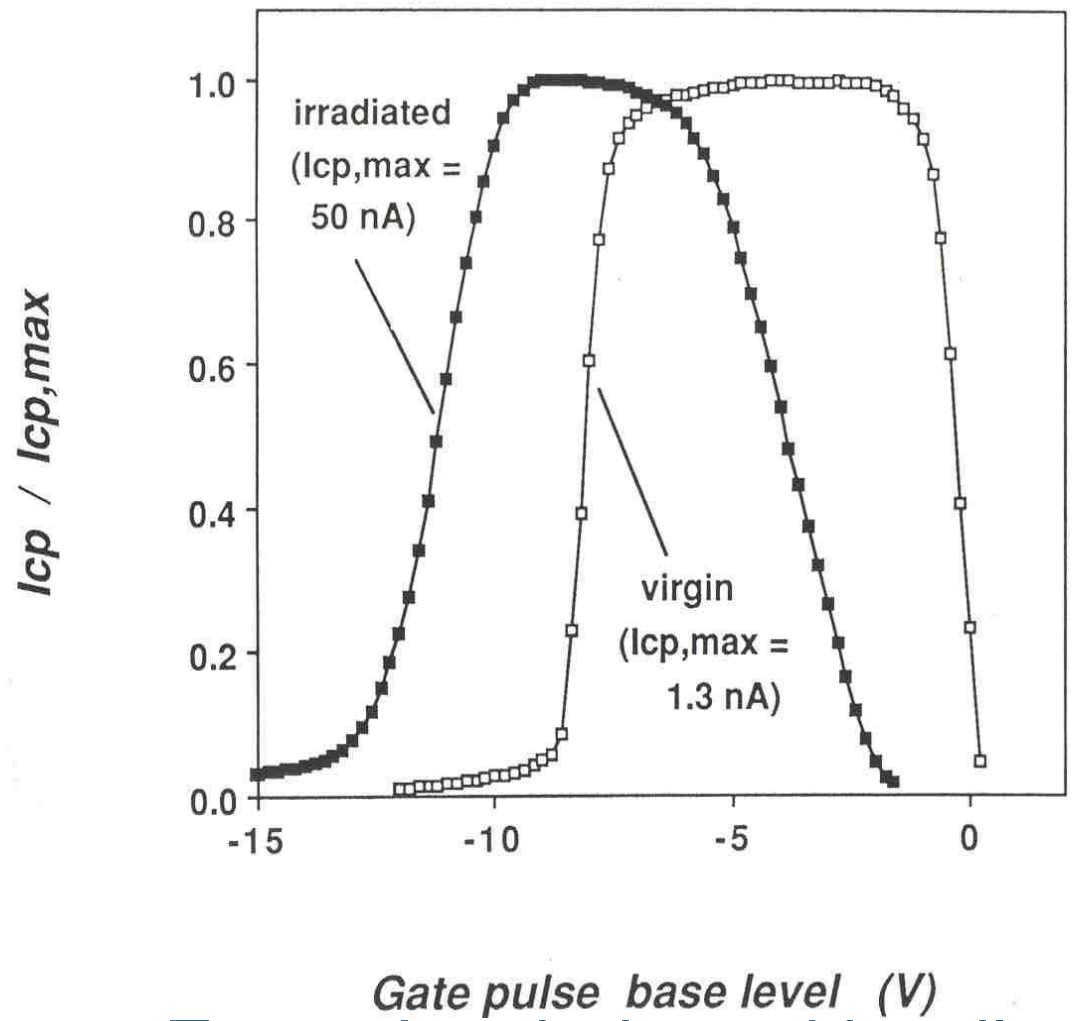
$D_{it} = 10^{10}$ $cm^{-2}eV^{-1}$ => spread = 10 mV

$D_{it} = 10^{11}$ $cm^{-2}eV^{-1}$ => spread = 100mV

$D_{it} = 10^{12}$ $cm^{-2}eV^{-1}$ => spread = 1 V

Interface trapped charge leads to a spread in the transition regions of the base level curves

IMPROVEMENTS TO THE MODEL



E = 25 keV

$Q_{e\text{-}beam}$ = 5x10^{-5} C/cm^2

$I_{e\text{-}beam}$ = 0.5x10^{-7} A

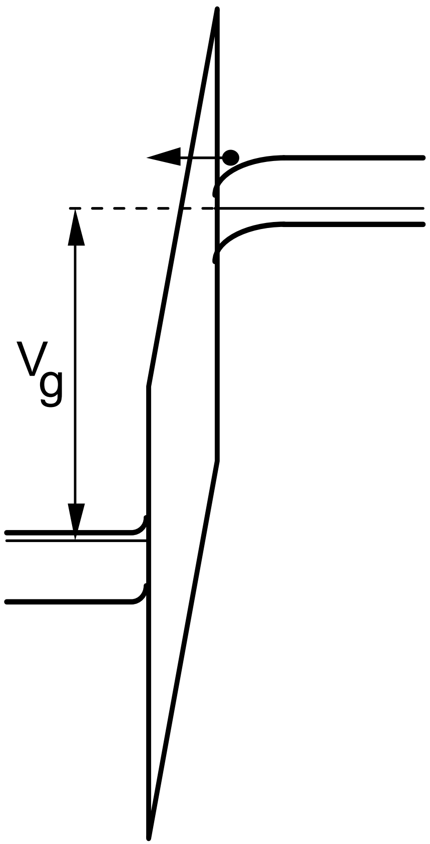
time = 250s

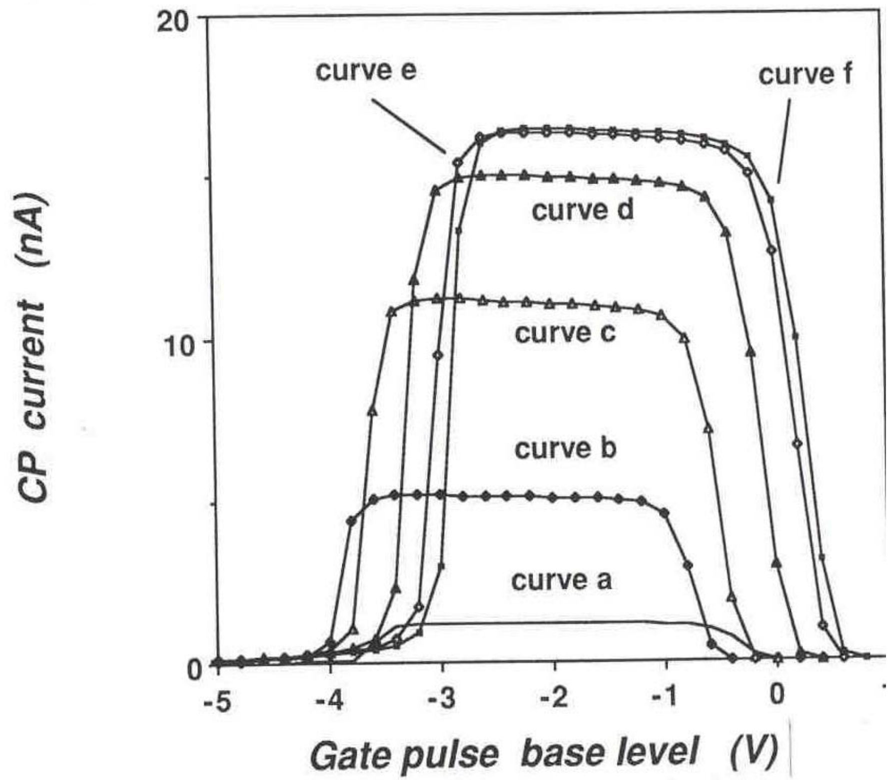
Example: virgin and irradiated MOSFET

Outline

1. Introduction
2. Basic CP-principle
3. Second order model
4. Temperature dependence
5. Base level edges
6. MOSFET degradation
7. Energy distribution
8. Lateral en vertical profiling
9. Geometric components
10. Effects of oxide thickness scaling
11. Single trap characterization
12. CP in high k gate stacks

Application: MOSFET degradation



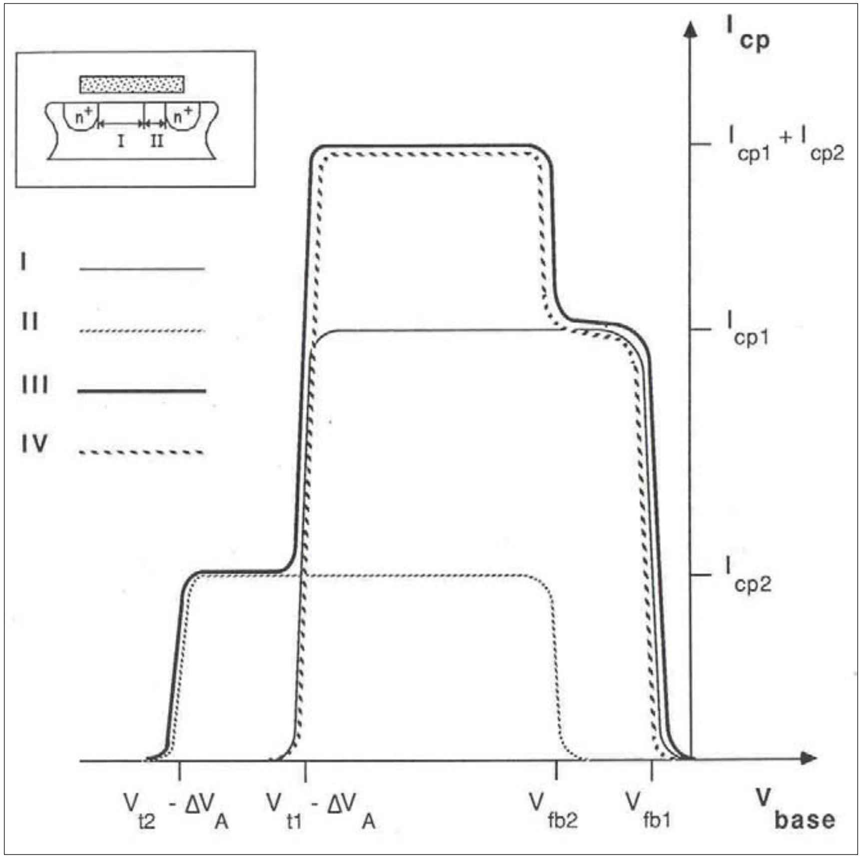


Vg = +12V, t = 0s (a), 1s (b), 10s (c), 100s (d), 500s (e)
Vg = -12V, t=1s (f)

(Heremans et al., IEEE TED, p. 1318, 1989)

Change in CP-curves under high-field Fowler-Nordheim injection

Sensitive to non-uniform degradation



Curve I :
CP of region I

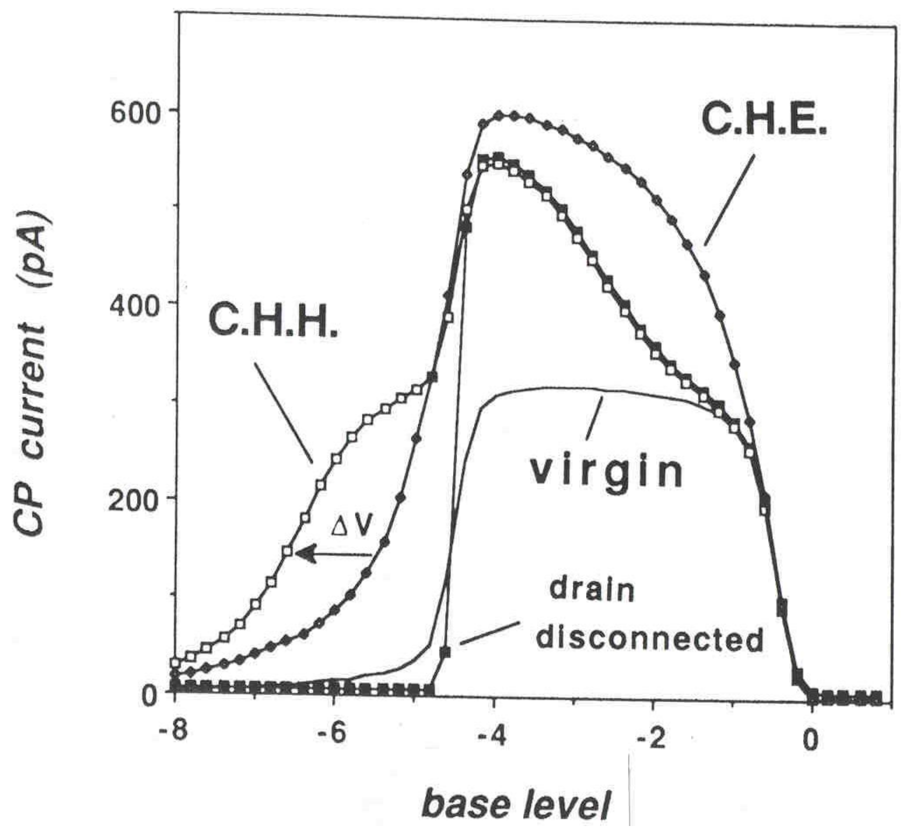
Curve II :
CP of region II

Curve III :
CP of whole transistor

Curve IV :
CP of whole transistor with disconnected drain

Example: positive charge and interface traps in n-channel MOSFET's

Non-uniform degradation of MOSFET's



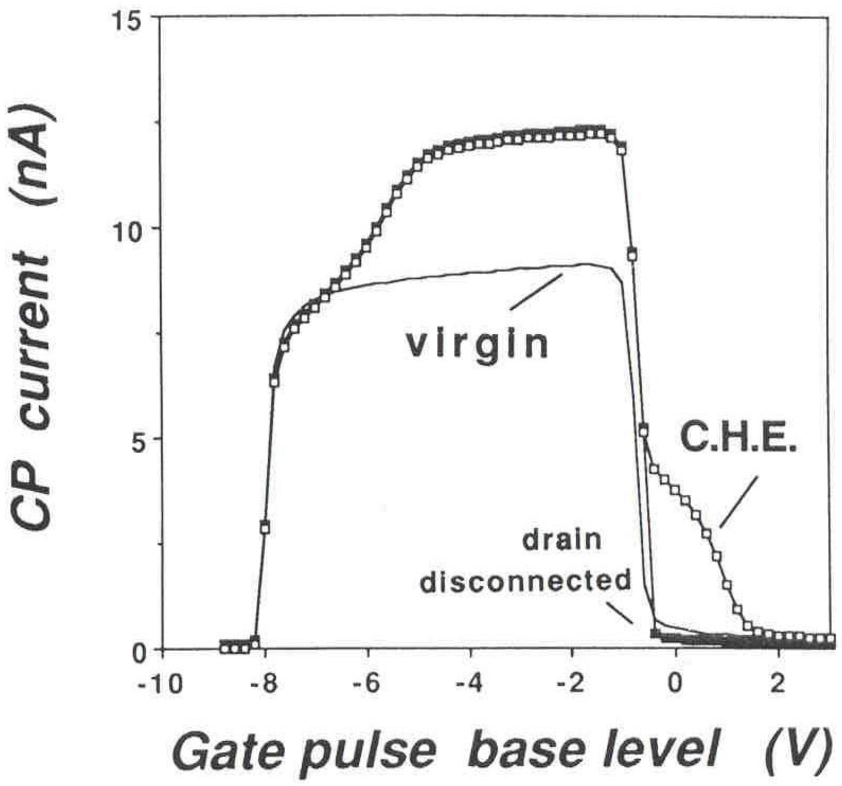
t_{ox} = 27nm
L_{eff} = 1.7μm
W = 100 μm

Stress conditions:

V_g = 1V
V_{ds} = 8V
t_{stress} = 1000s

Example: channel hot hole injection in n-MOSFET

NON-UNIFORM DEGRADATION OF MOSFET's



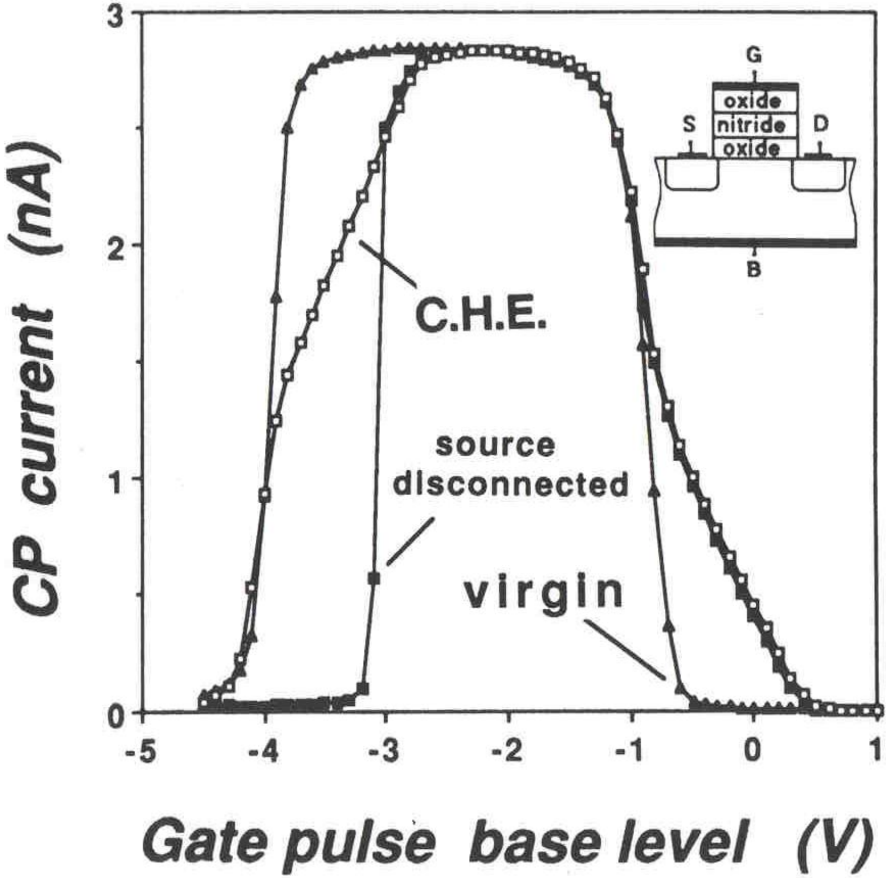
t_{ox} = 28nm
L_{eff} = 1.7μm
W = 200 μm

Stress conditions:

V_g = -1.4V
V_{ds} = -8.5V
t_{stress} = 1000s

Example: channel hot hole injection in p-MOSFET

NON-UNIFORM DEGRADATION OF MOSFET's



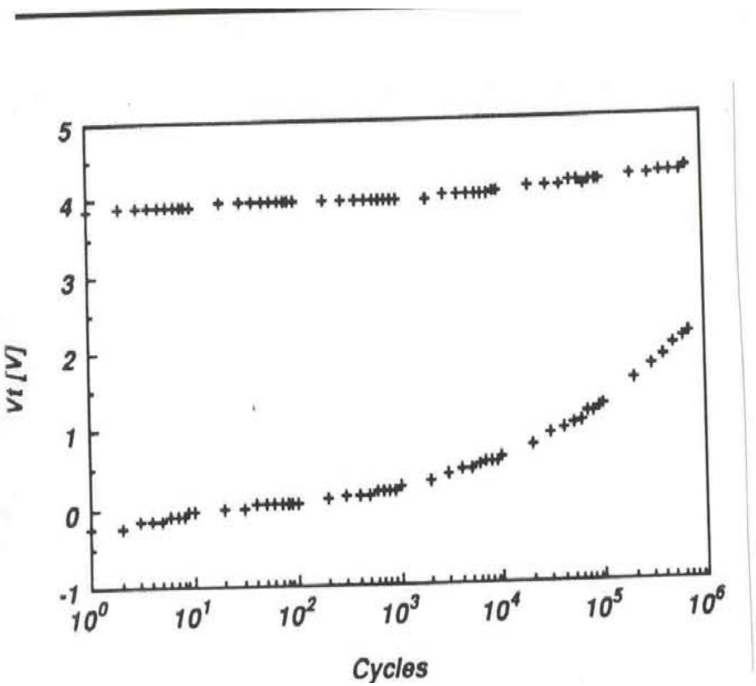
$t_{ox}1 = 2$ nm
$t_{nitride} = 30$ nm
$t_{ox}2 = 5$ nm
$L_{eff} = 6$ µm
$W = 20$ µm

Stress conditions:

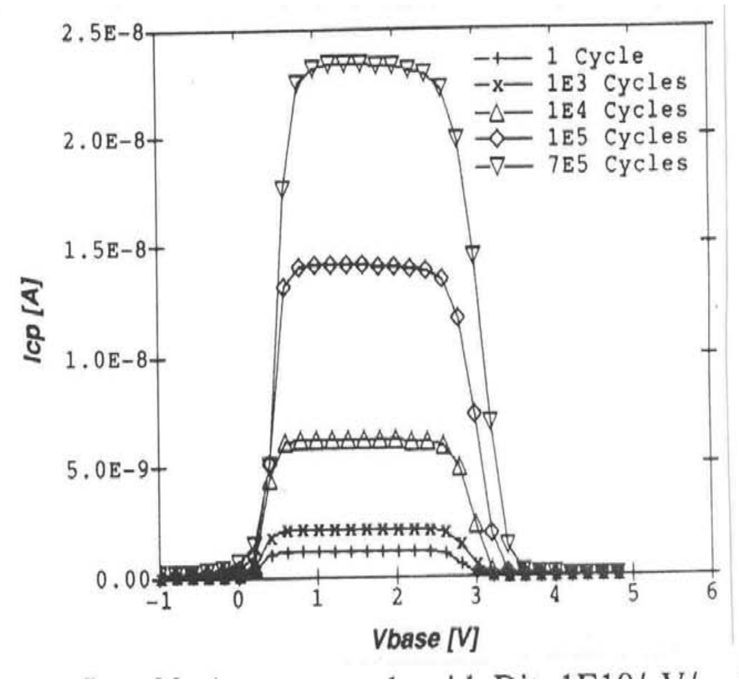
$V_g = 12V$
$V_{ds} = 12V$
$t_{stress} = 10s$

Example: channel hot electron injection in an n-SONOS transistor

NON-VOLATILE MEMORY CELL DEGRADATION



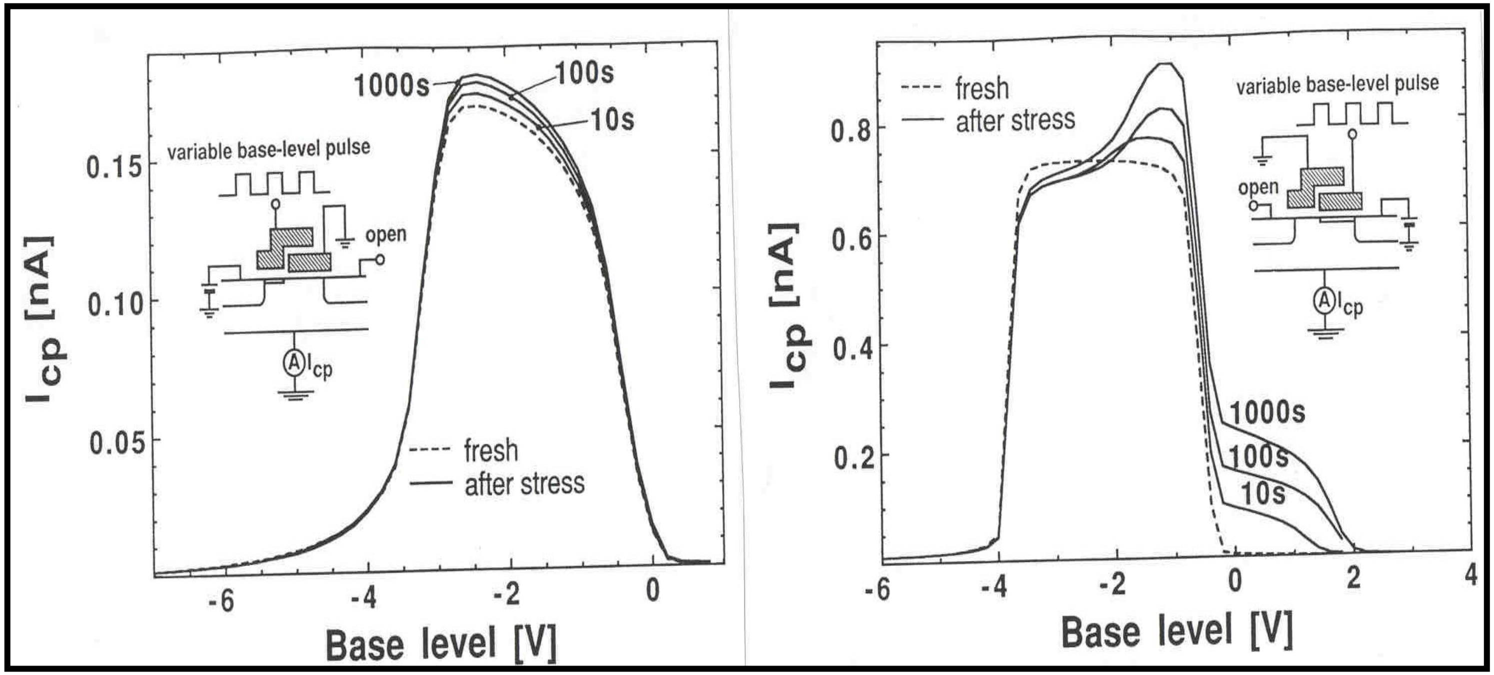
Endurane characteristic for uniform WRITE operation



Charge pumping characteristics for WRITE operation

Floating gate memory cells

D. Wellekens et al, IEEE TED, p. 1992, 1995

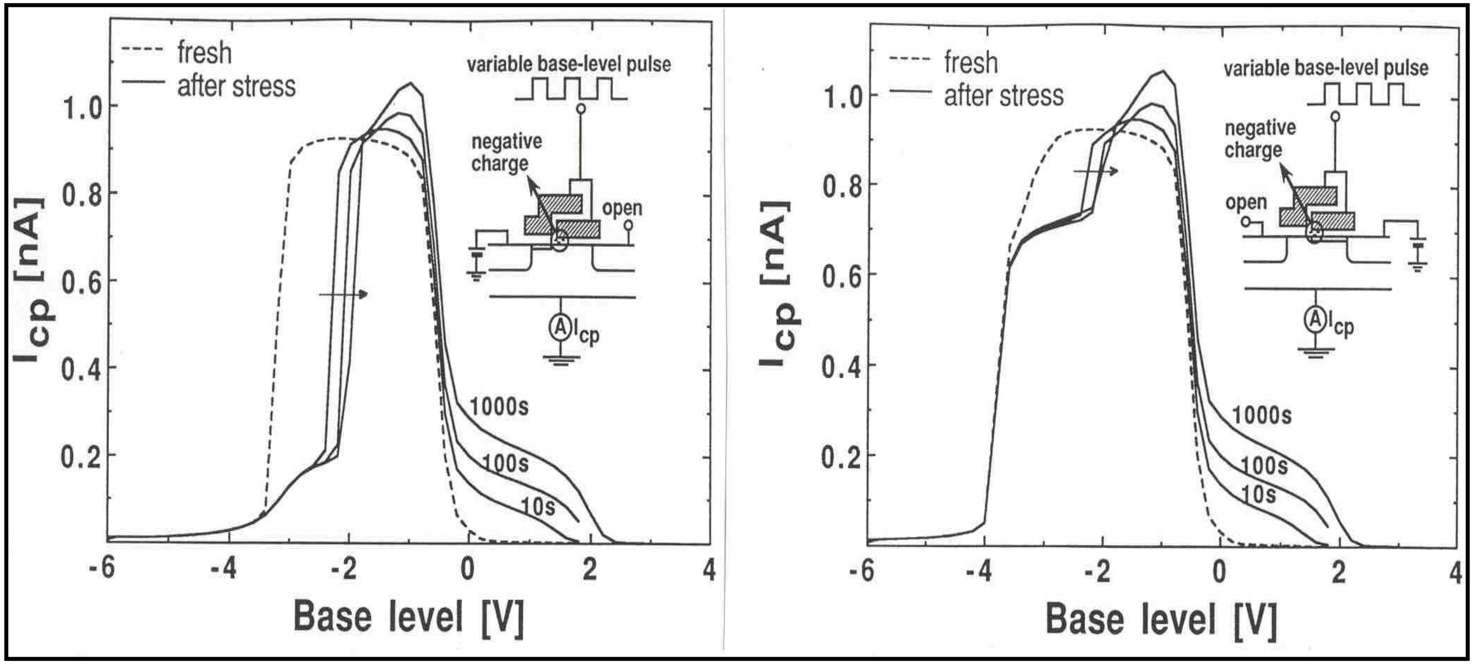


Source gate area:
no degradation observed

Drain gate area:
interface traps and negative trapped charge observed

Degradation characteristics of split-gate transistors

D. Wellekens et al, IEEE TED, p. 1992, 1995



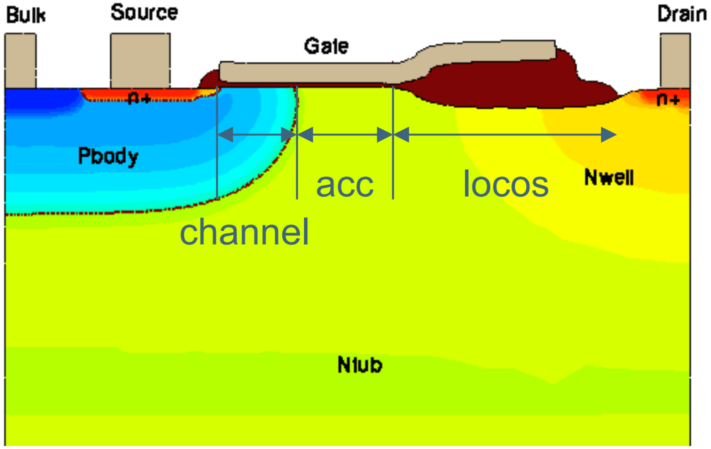
Full transistor pumping with drain disconnected

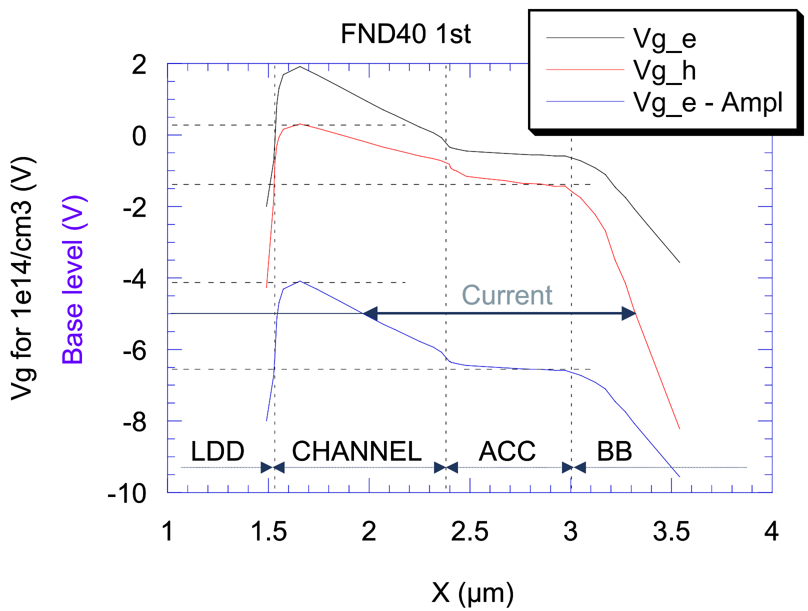
Full transistor pumping with source disconnected

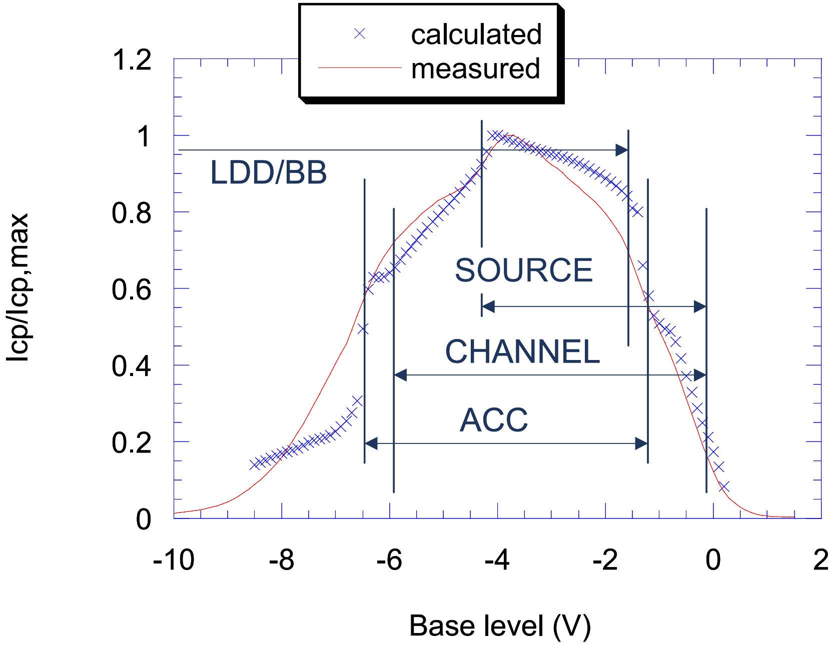
Degradation characteristics of split-gate transistors

Charge Pumping on LDMOS devices

- 40V LDMOS
- V_{ge} and V_{gh} as from TCAD
- Uniform N_{it} in thin oxide



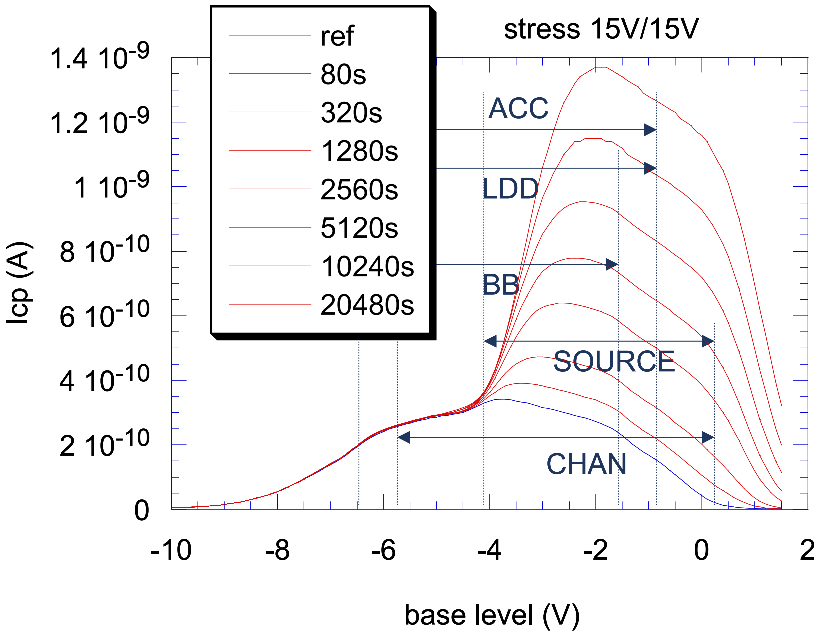




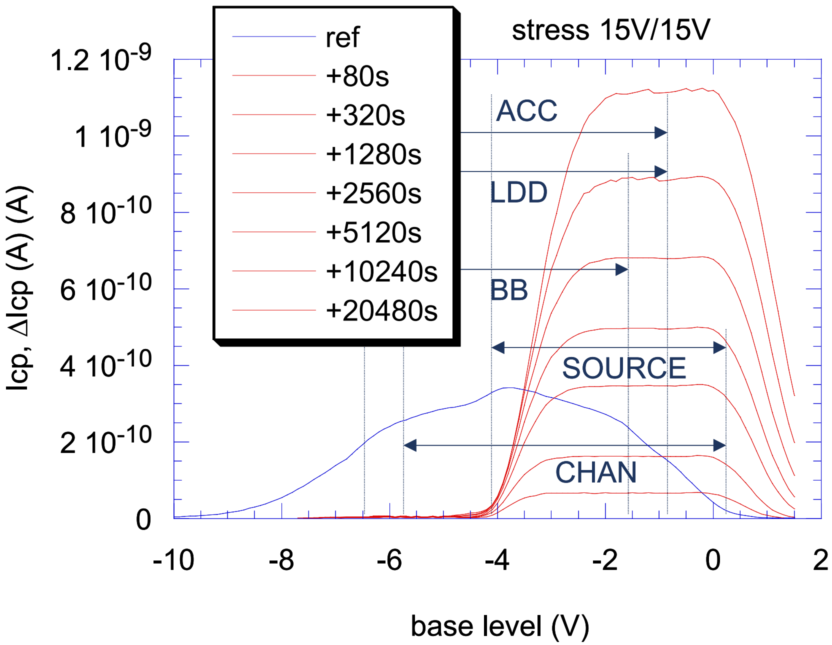
Charge Pumping on LDMOS devices

- LDMOS stressed at $V_{ds}=V_{gs}=15V$ (low V_{ds}, high V_{gs}) → N_{it} formation at the source.

Charge pumping spectra



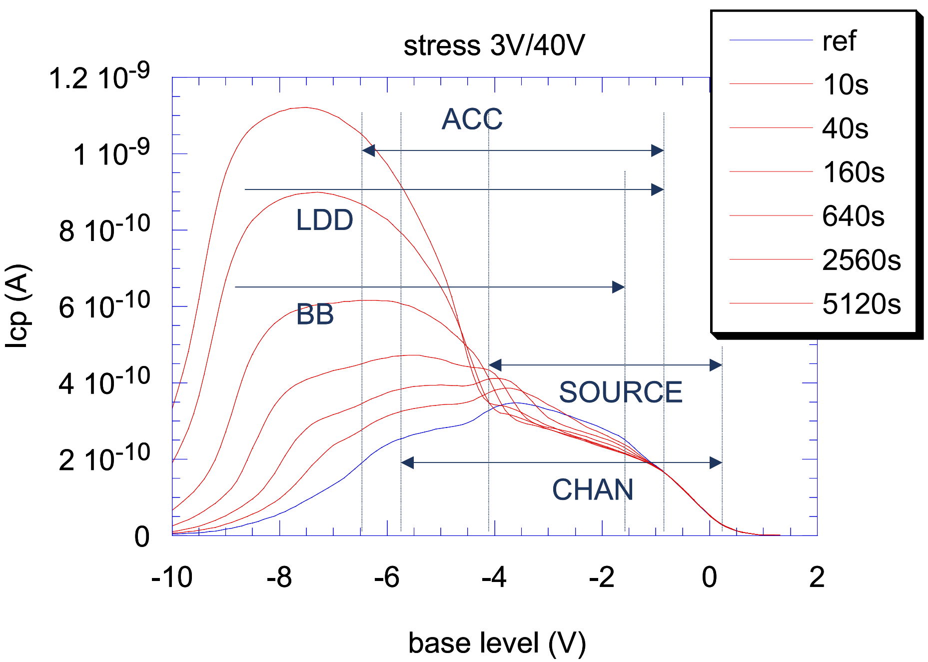
Differential CP spectra

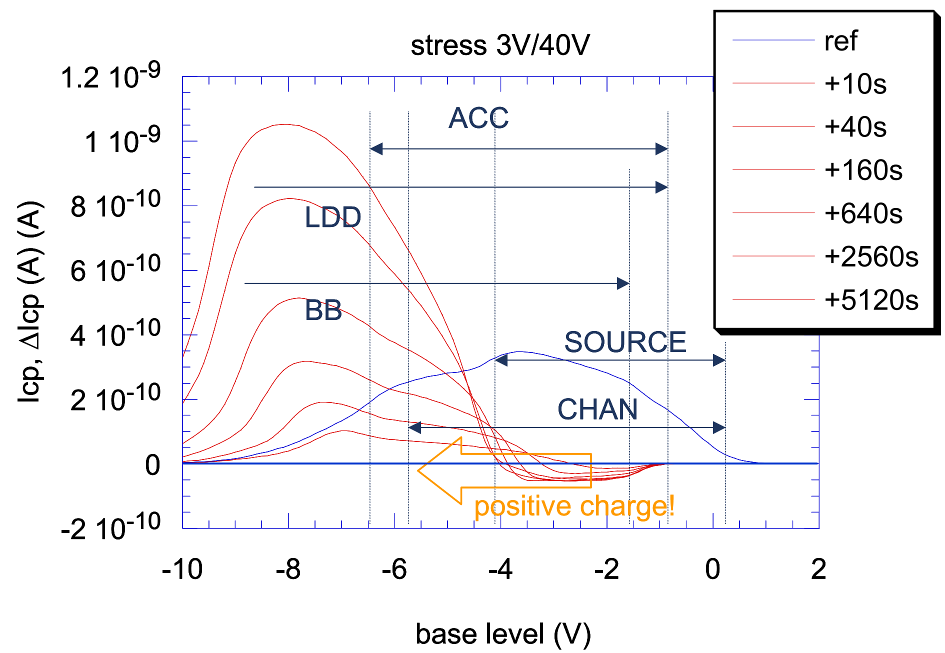


P. Moens et al, IRPS Tutorial 2005

Charge Pumping on LDMOS devices

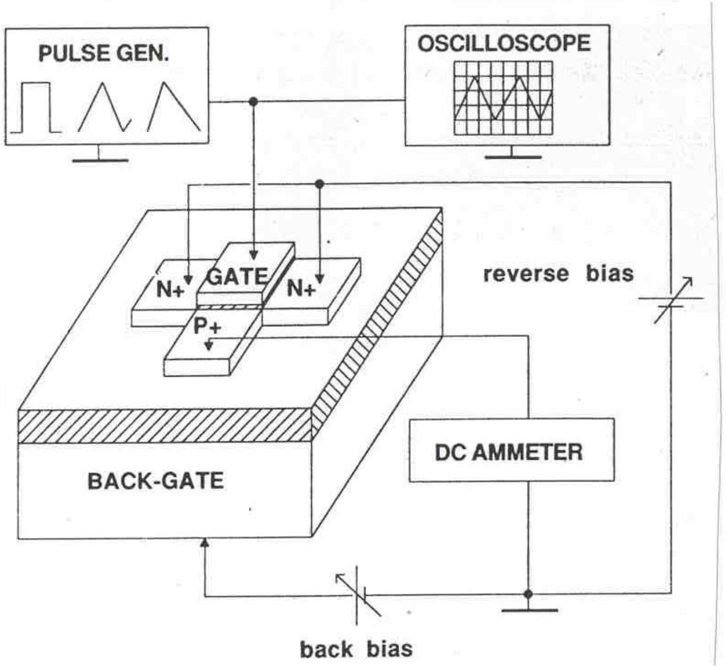
- LDMOS stressed at V_{ds}=40 V_{gs}=3V (low V_{gs}, high V_{ds}) → N_{it} formation in accumulation region or under the birds beak.

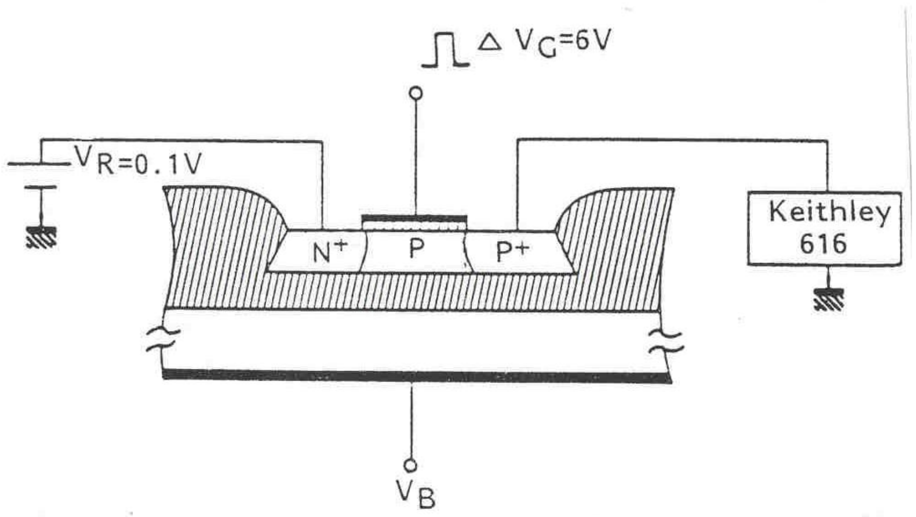




P. Moens et al, IRPS Tutorial 2005

SOI-CHARACTERIZATION





Five terminal SOI-MOSFET

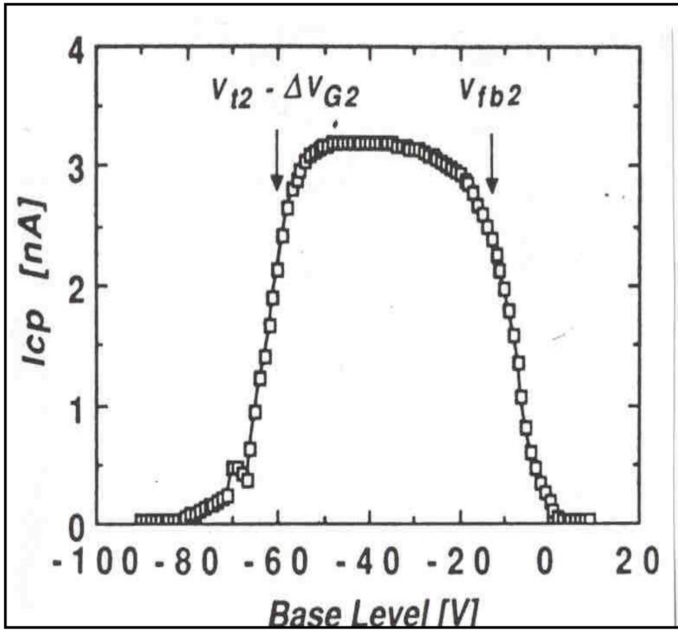
[Wouters et al, 1989]

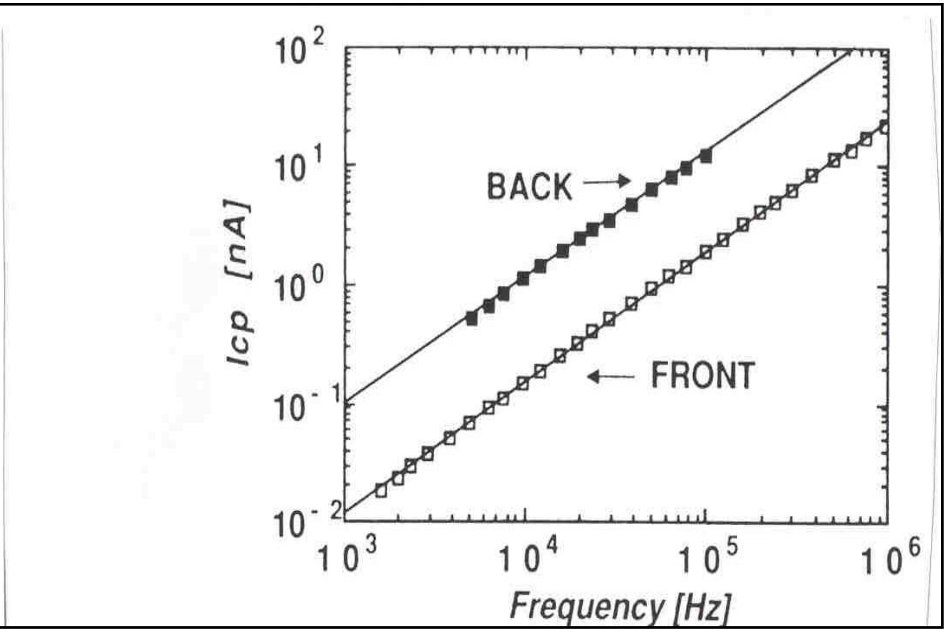
Gated P-I-N diode

[Elewa et al, 1988]

SOI-MOSFET Structures for charge pumping

SOI-CHARACTERIZATION



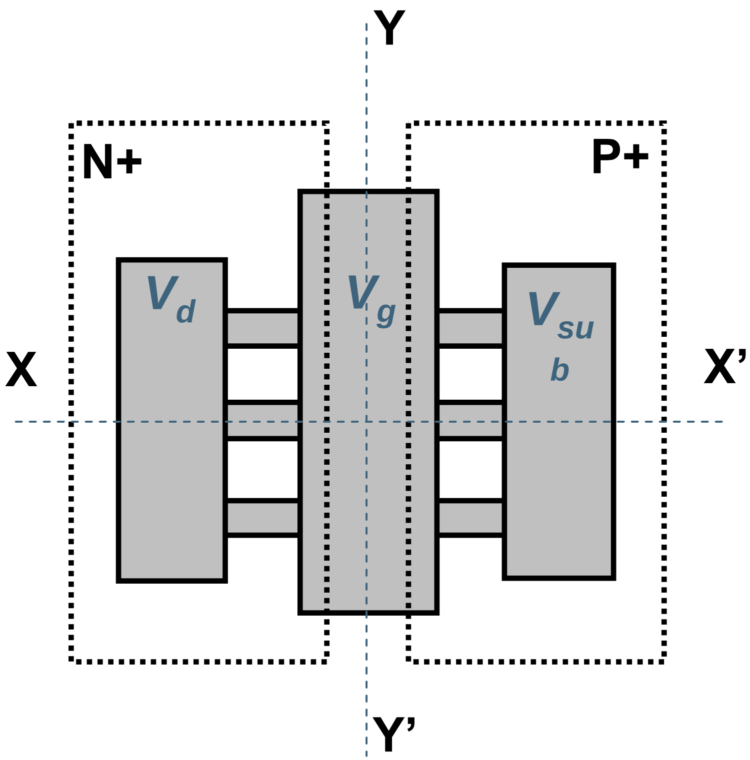


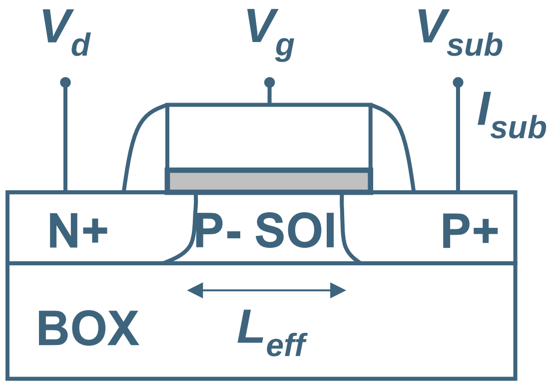
Base level curve of the back interface

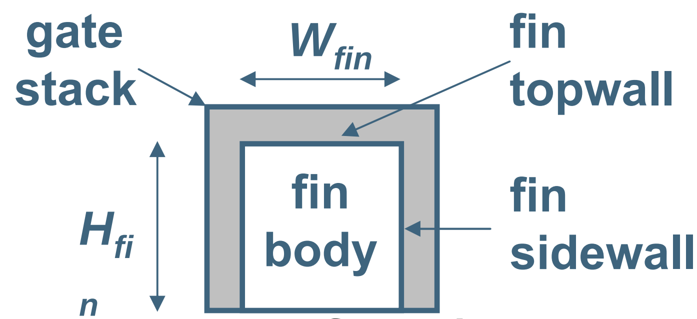
Frequency dependence of Icp for back and front interface

- SOI-MOSFET charge pumping characteristics

Characterization of Finfet interfaces



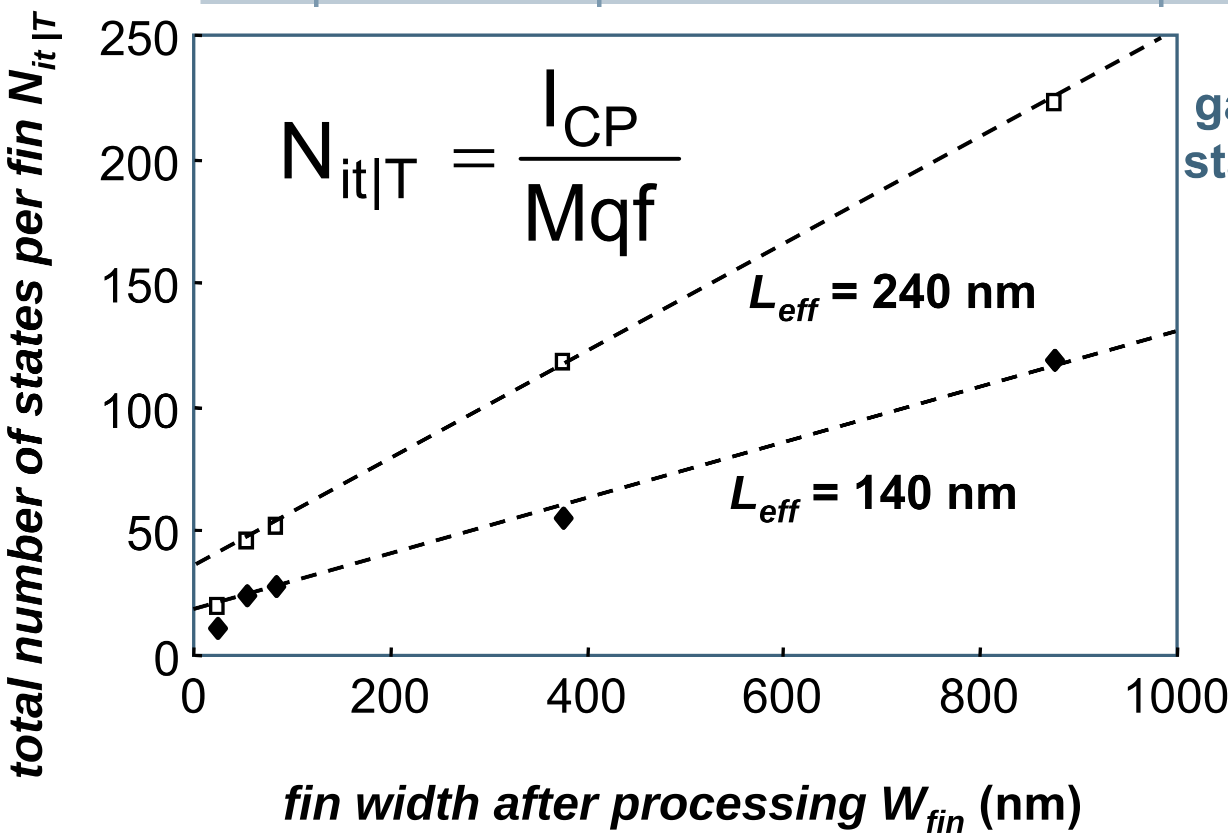


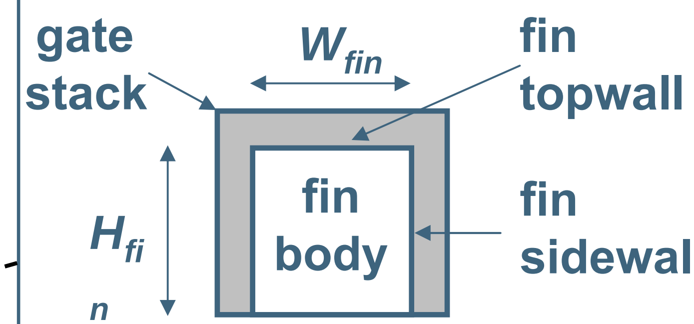


- Sidewall interface quality important for drive current !
- Use charge pumping on gated Fin-diode with various geometries

Characterization of fin interfaces

$$N_{it|T} = 2N_{it|SW} L_{eff} H_{fin} + N_{it|TW} L_{eff} W_{fin}$$





$$N_{it|SW} = \frac{c}{2L_{eff} H_{fin}}$$

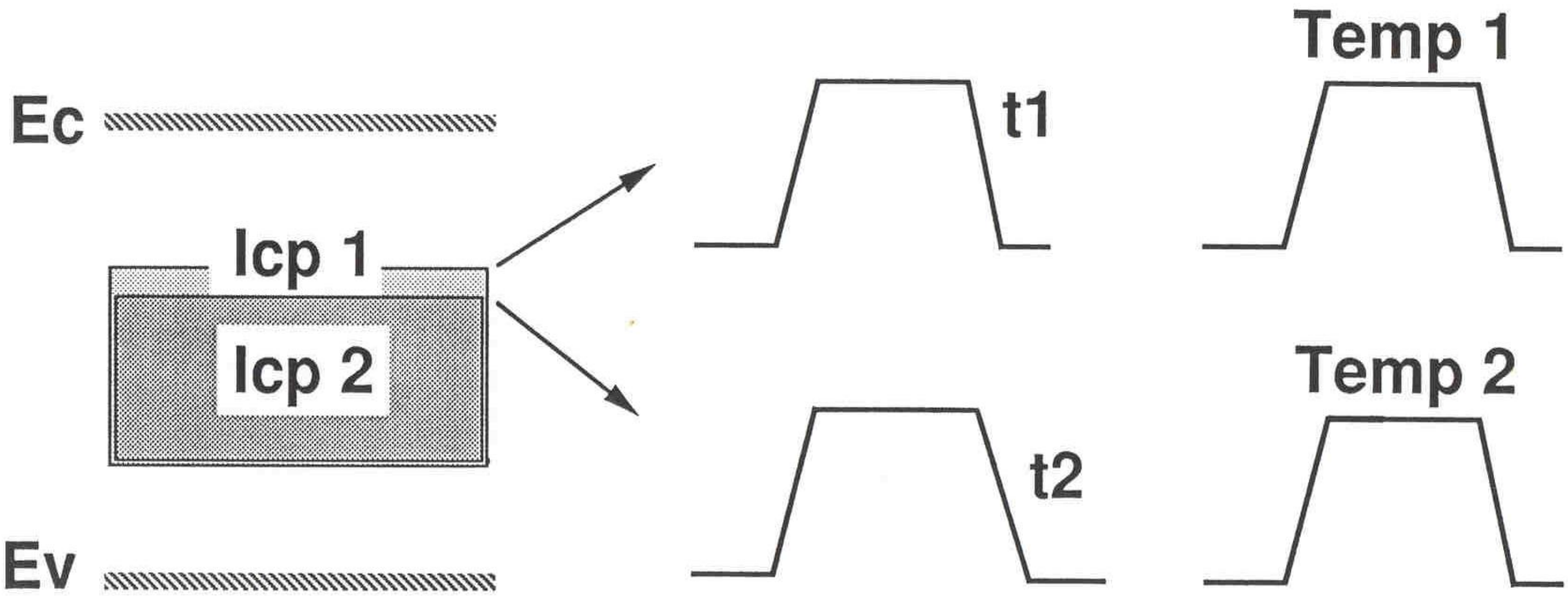
$$N_{it|TW} = \frac{m}{L_{eff}}$$

Kapila et al., IEEE Electron Dev. Lett., vol. 28, p. 232, 2007

Outline

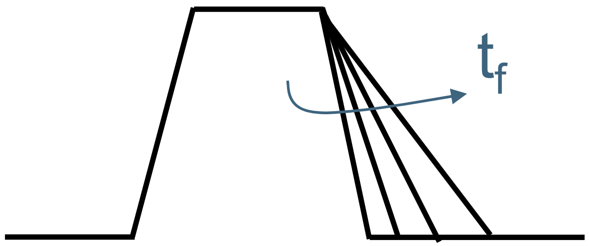
1. Introduction
2. Basic CP-principle
3. Second order model
4. Temperature dependence
5. Base level edges
6. MOSFET degradation
7. Energy distribution
8. Lateral and vertical profiling
9. Geometric components
10. Effects of oxide thickness scaling
11. Single trap characterization
12. CP in high k gate stacks

Energy profiling



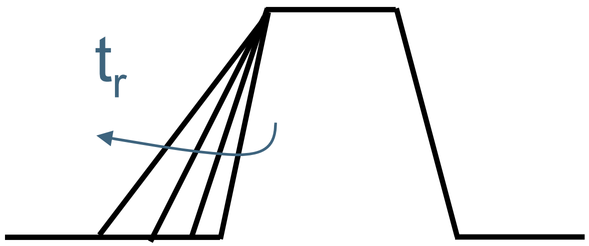
Emission levels can be modulated
by rise/fall times or by temperature

Energy profiling: method 1



changing t_f → scan in top part of the bandgap

$E_{em,e}(t) = E_i + kT\ \ln(\sigma_n\ v_{th}\ n_i\ t_{em,e})$



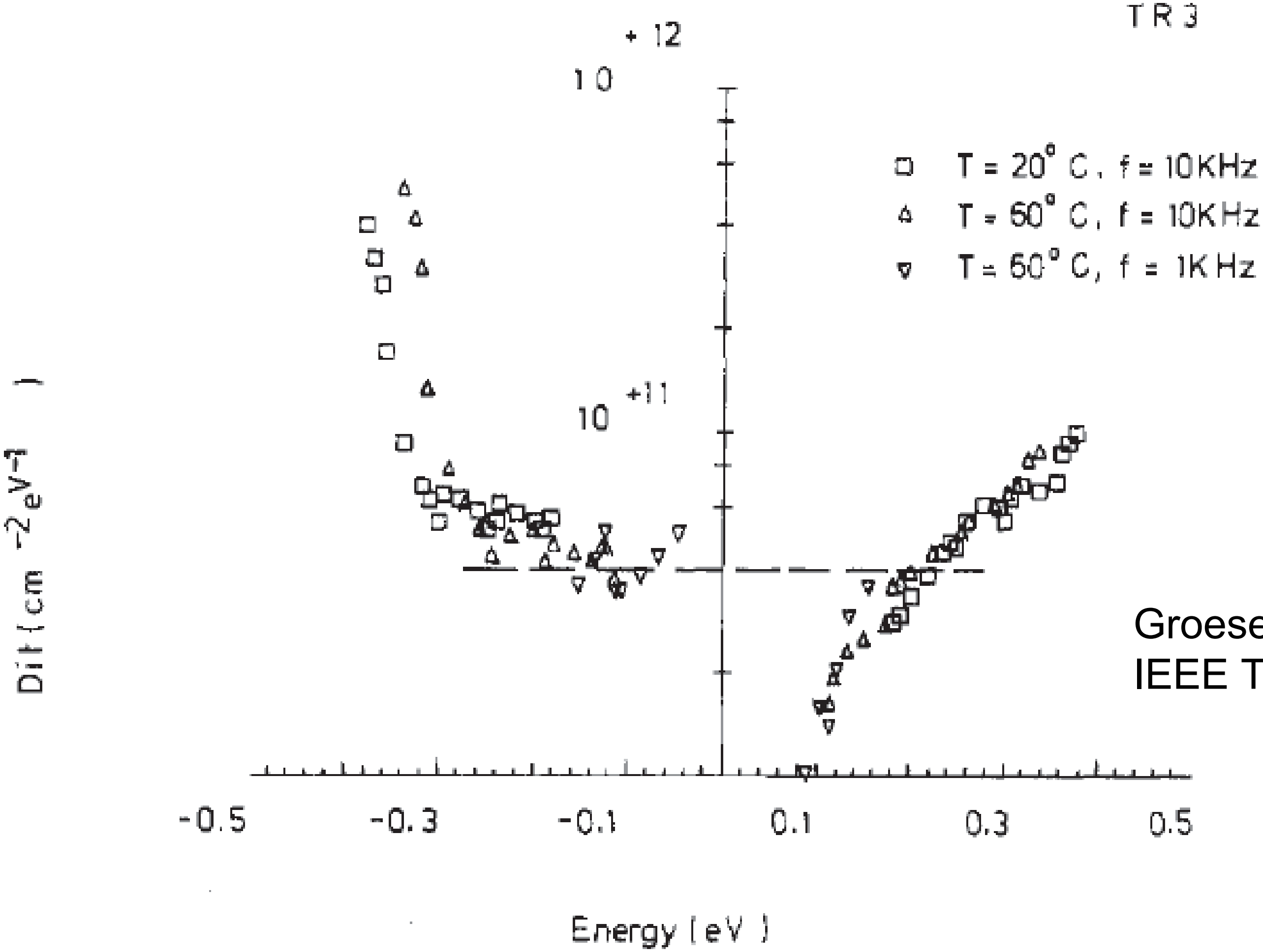
changing t_r → scan in bottom part of the bandgap

$E_{em,h}(t) = E_i + kT\ \ln(\sigma_p\ v_{th}\ n_i\ t_{em,e})$

$$D_{it}(E_{em,e}) = -\frac{1}{q \cdot A \cdot kT \cdot f}\frac{dI_{cp}}{d\ln t_f}$$

$$D_{it}(E_{em,h}) = -\frac{1}{q \cdot A \cdot kT \cdot f}\frac{dI_{cp}}{d\ln t_r}$$

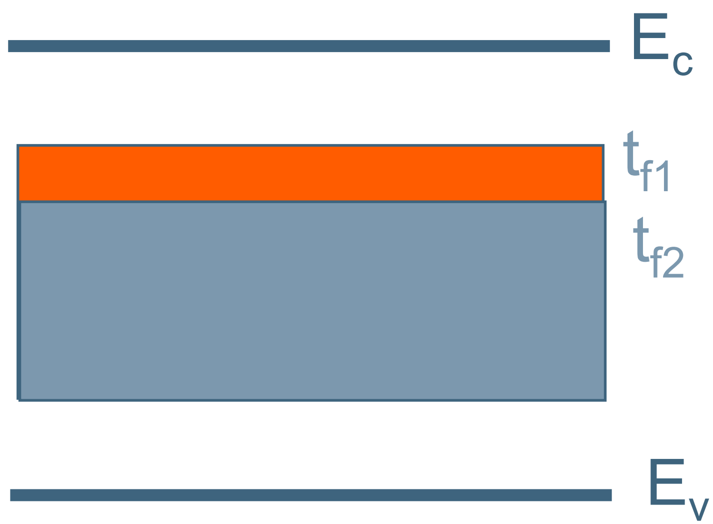
Energy profiling: method 1

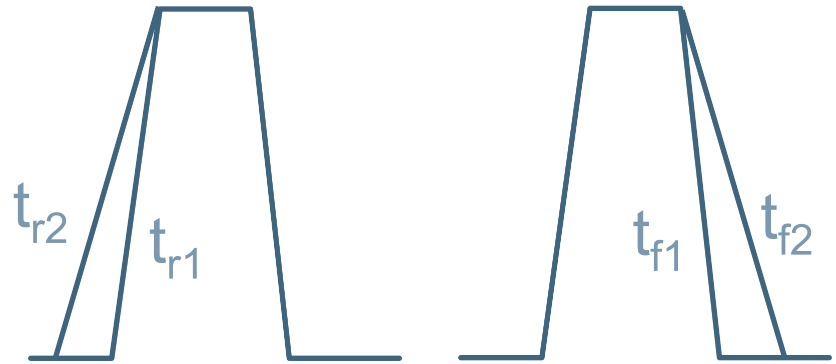


Groeseneken et al
IEEE TED, p. 42, 1984)

Example of energy distribution of interface traps using Method 1

Energy profiling: method 2





$$S_{f,r} = I_{cp}(t_{f,r\,1}) - I_{cp}(t_{f,r\,2}) = q \cdot A \cdot f \cdot D_{it}(E_o) \cdot \Delta E_{f,r}$$

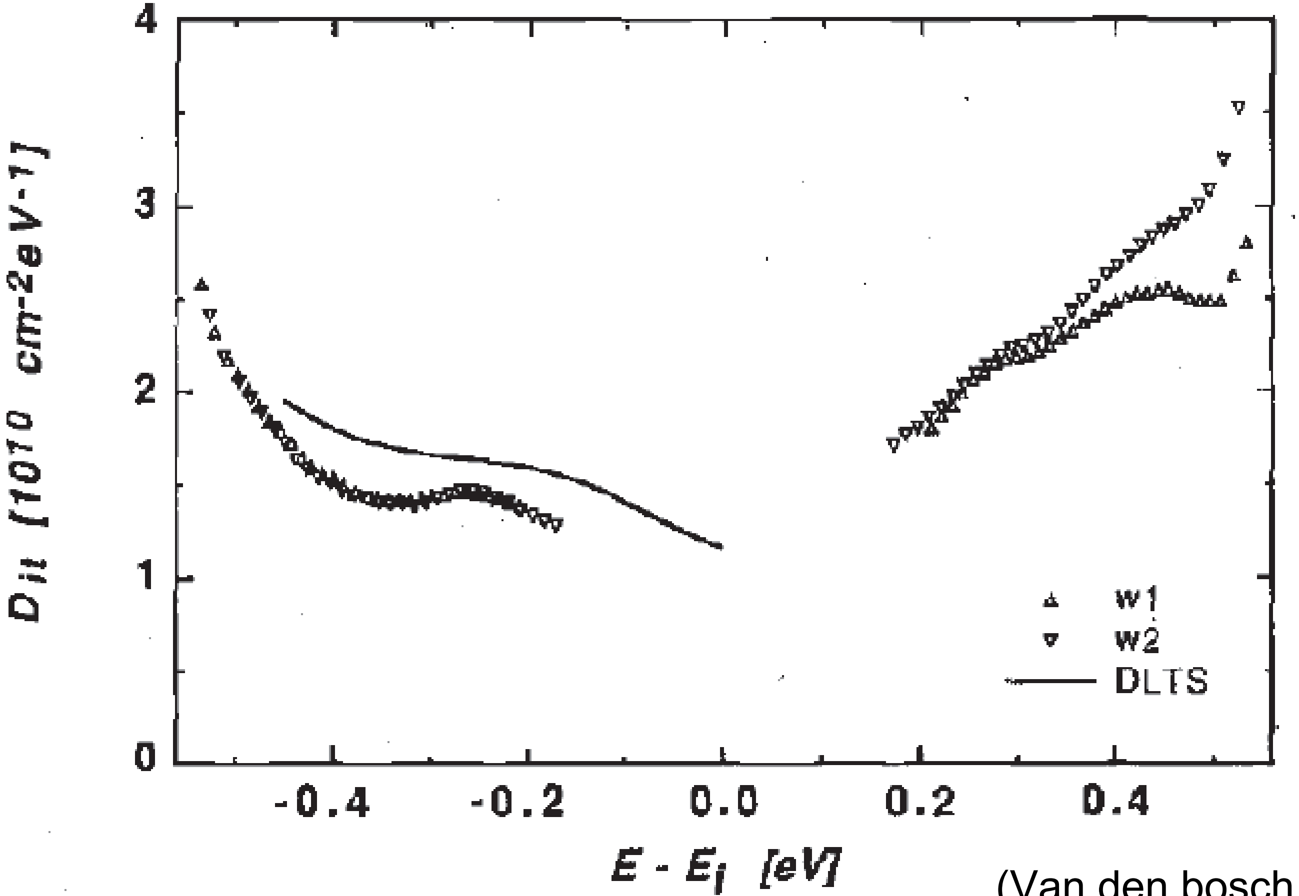
$$\Delta E_{f,r} = kT \cdot \ln\left(\frac{t_{f,r\,2}}{t_{f,r\,1}}\right)$$

$$D_{it}(E) = \frac{S_{f,r}(T)}{q \cdot f \cdot A \cdot \Delta E_{f,r}}$$

Varying temperature at fixed emission window ΔE

(Van den bosch et al. IEEE TED, p. 1820, 1991)

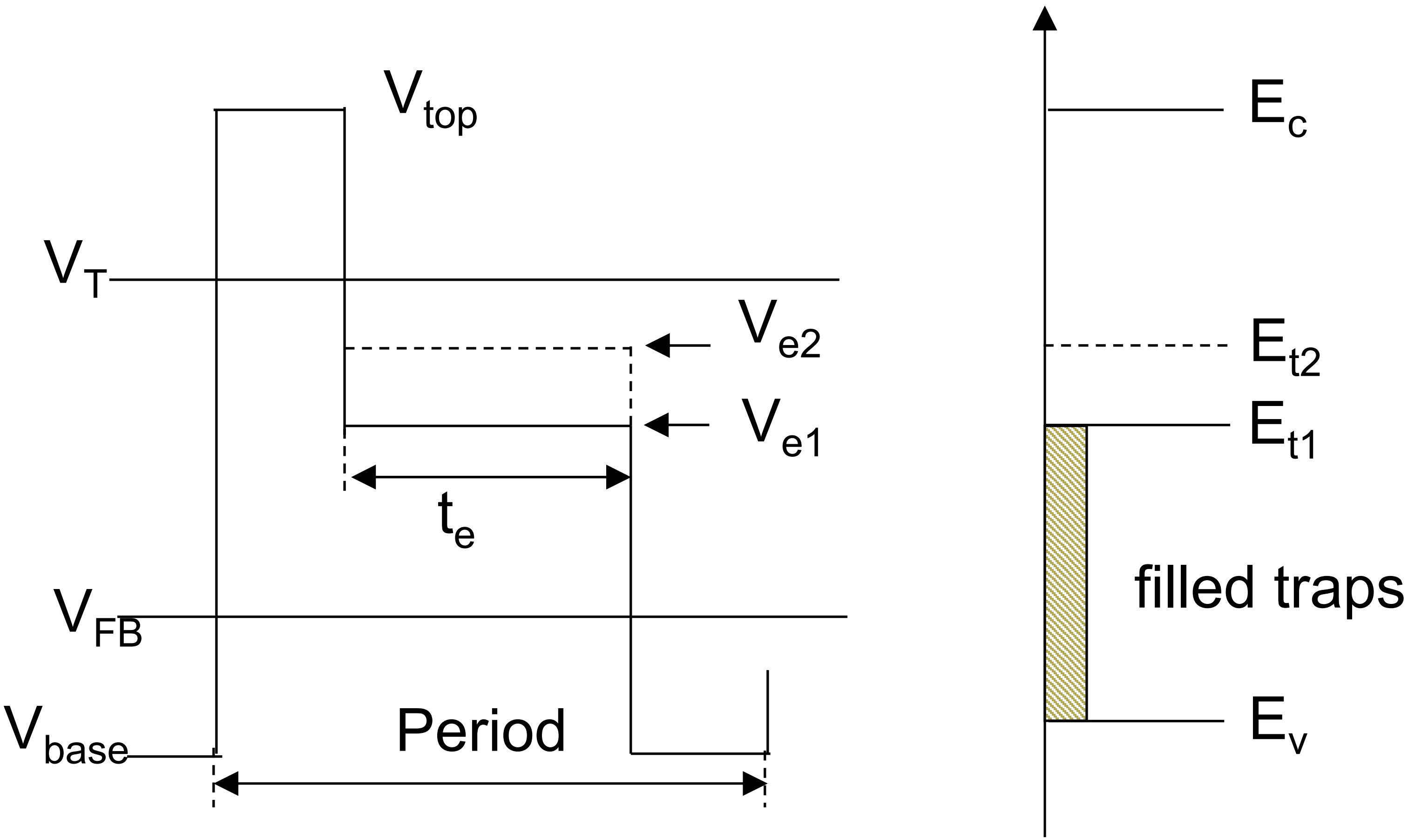
Energy profiling: method 2



(Van den bosch et al. IEEE TED, p. 1820, 1991)

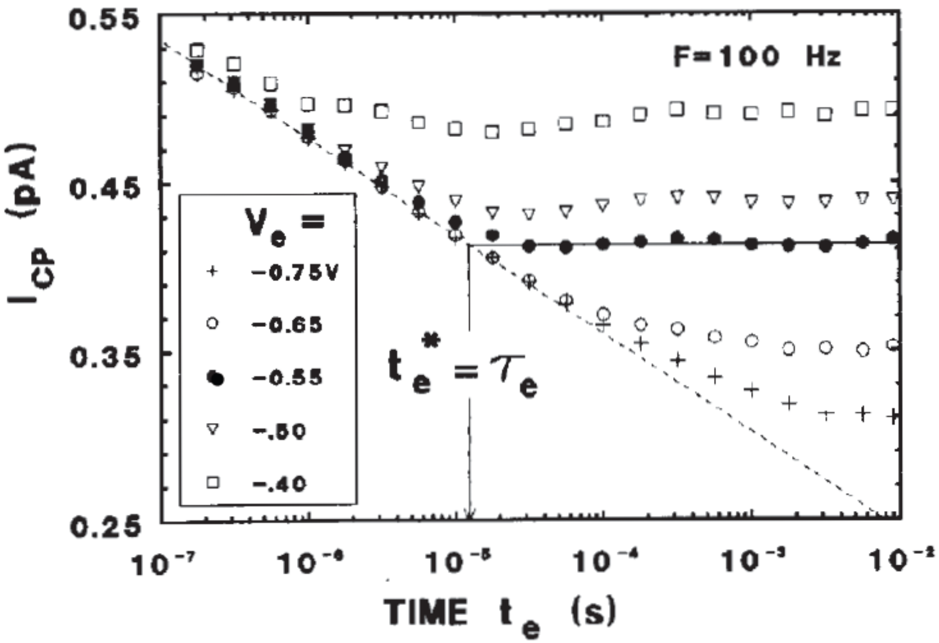
Example of energy distribution of interface traps using Method 2 (Spectroscopic charge pumping)

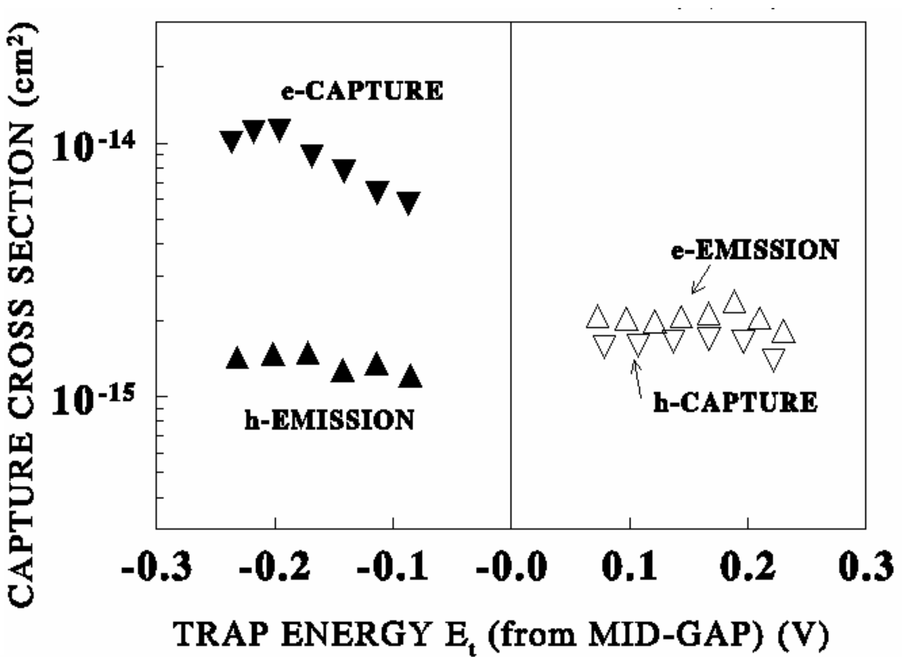
Energy profiling: method 3



$$D_{it}(E_t) = \frac{1}{q\,f\,A_G}\frac{dI_{cp}}{dV_e}\frac{dV_e}{d\psi_e}$$

Energy profiling: method 3





Icp vs length at midlevel t_e

(Ancona and Saks, J. Appl. Phys., p. 4415, 1992)

Energy distributions of electron and hole emission and capture cross sections

INTERFACE TRAP ENERGY DISTRIBUTIONS

Energy range:

Bandgap can be accessed from ± 0.52 eV to ±0.15 eV

- minimum midgap value is limited by diode reverse leakage current at high T, gate leakage current and lowest allowable measurement frequency
- midgap region is addressed by the variable base level CP-technique

Sensitivity:

Measurement sensitivity is in the range of 10^9 - 10^{10} cm^{-2} eV^{-1} depending on transistor size and resolution of the current meter

- in the integral form, the sensitivity of CP can be as high as 10^8 cm-2 eV-1, which is two order of magnitudes better than other interface characterization techniques (CV, conductance)

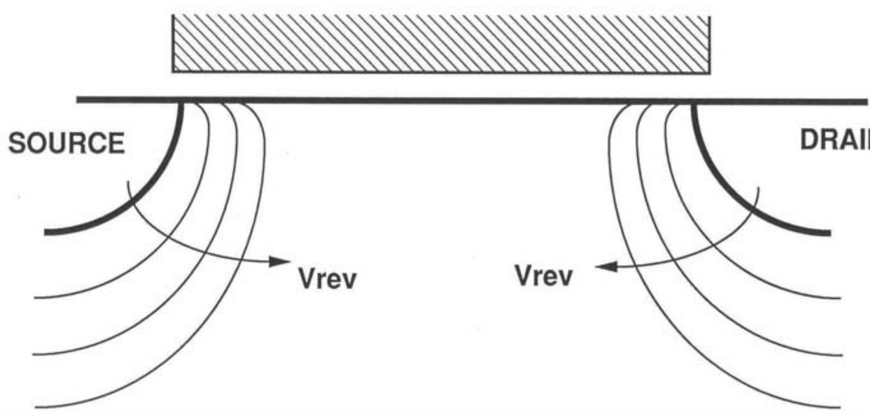
Energy resolution:

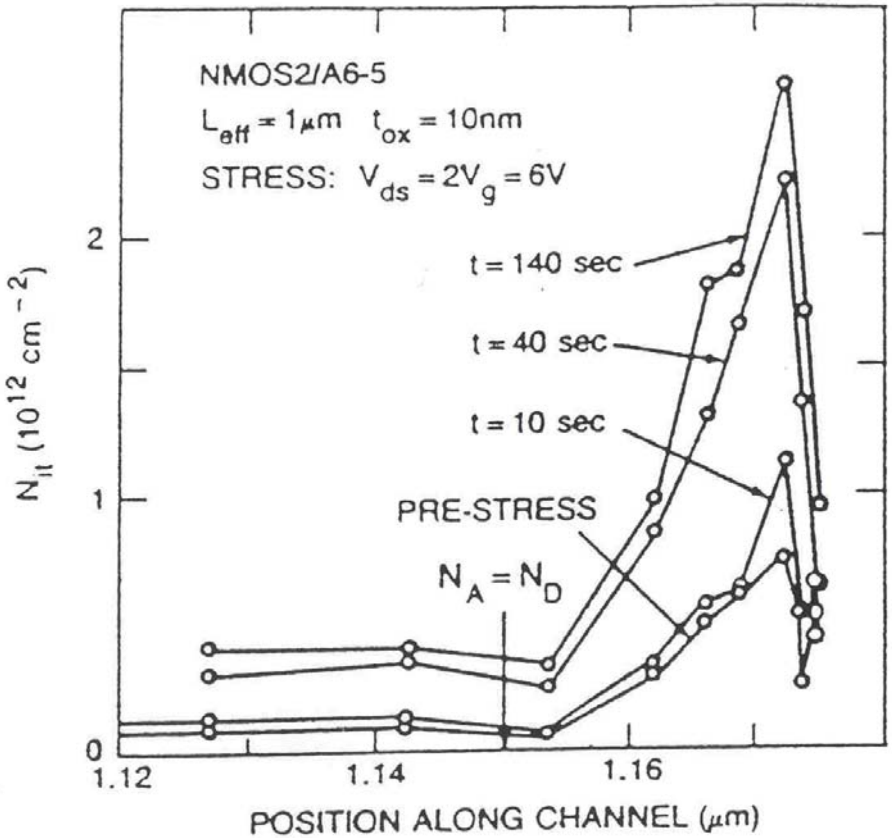
Energy resolution is in the order of kT, and thus improves at lower temperatures

Outline

1. Introduction
2. Basic CP-principle
3. Second order model
4. Temperature dependence
5. Base level edges
6. MOSFET degradation
7. Energy distribution
8. Lateral and vertical profiling
9. Geometric components
10. Effects of oxide thickness scaling
11. Single trap characterization
12. CP in high k gate stacks

Lateral profiling: method 1





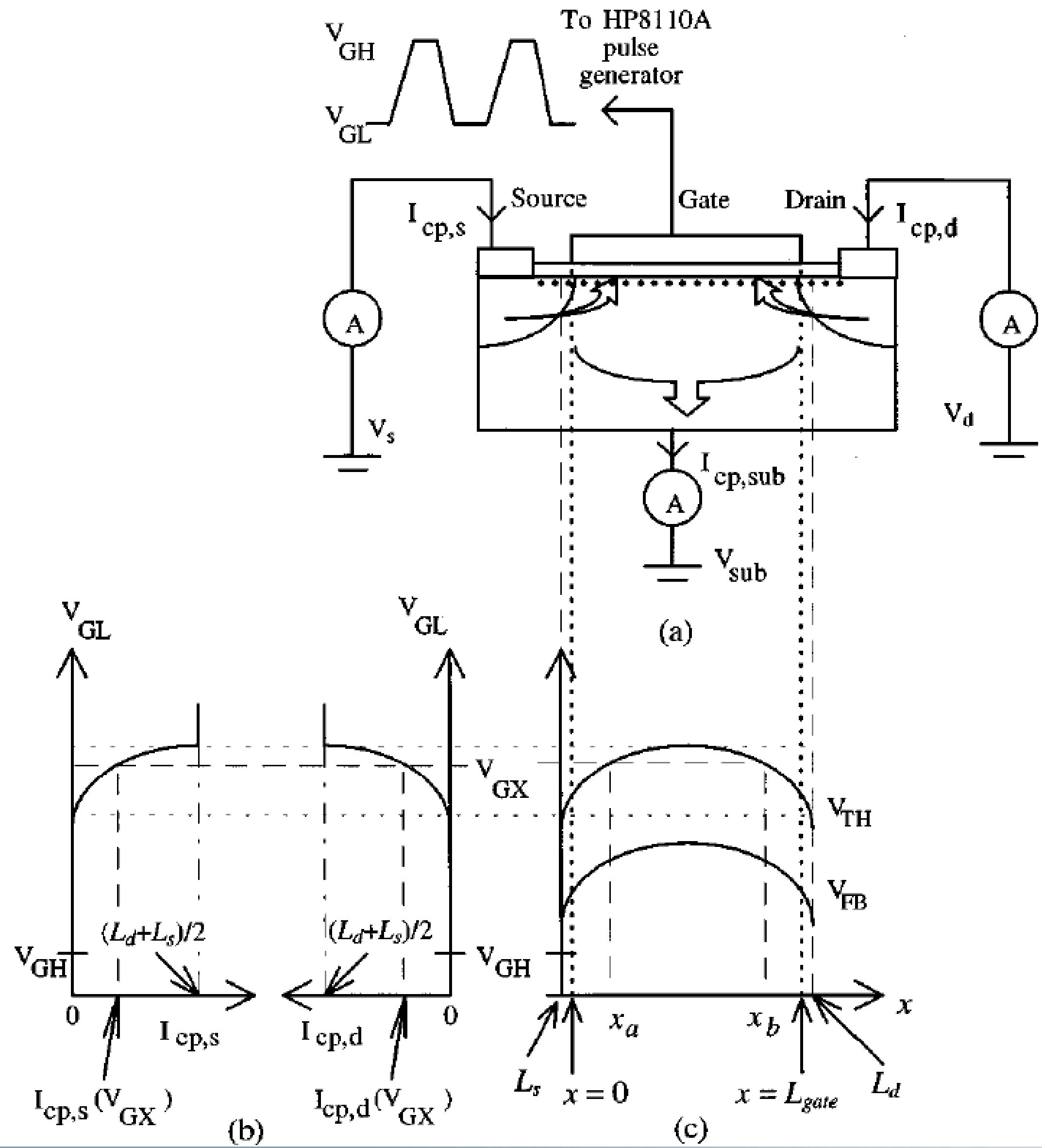
Ancona & Saks, IEEE TED, p. 2221, 1988

First class of methods: scan the lateral distance by increasing the space charge region around source and drain (increase Vr

Lateral profiling:method 2

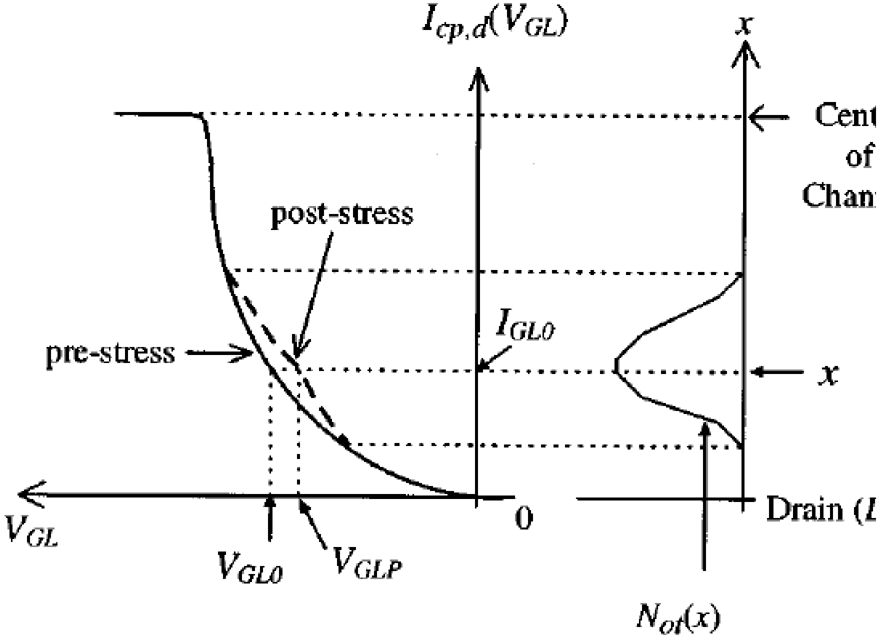
Second class of methods:

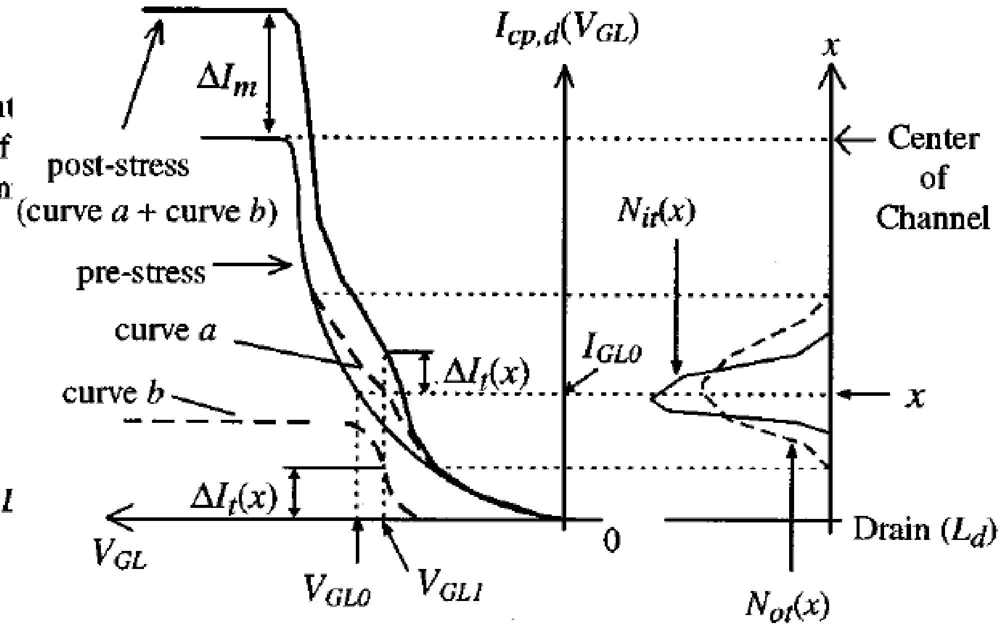
Get information on lateral profile from change of transition edges of base level or amplitude scans



Tsuchiaki et al., IEEE TED, p. 1768, 1993
Chim et al., J. Appl. Phys., vol. 81, p. 1992, 1997
Furnemont et al., IEEE EDL, p. 276, 2007

Lateral profiling: method 2

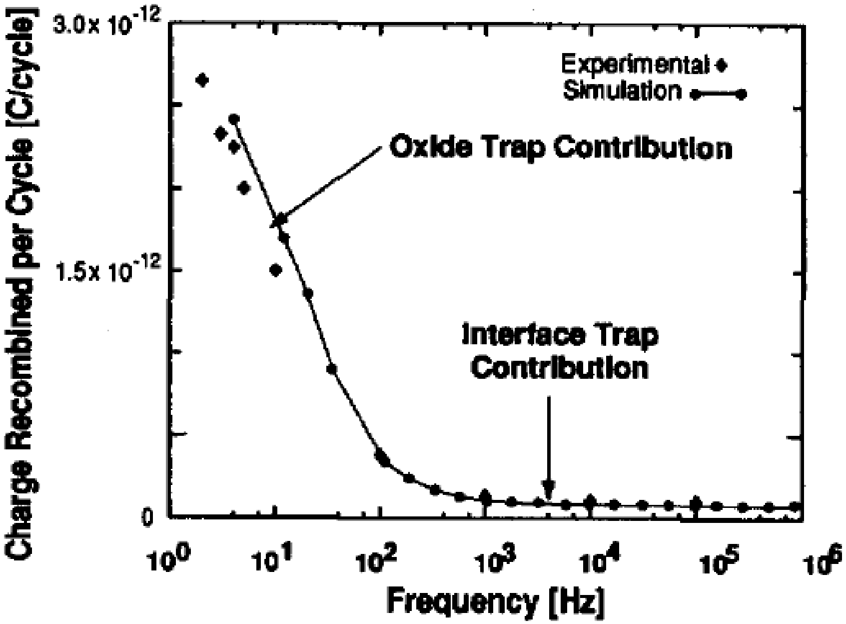


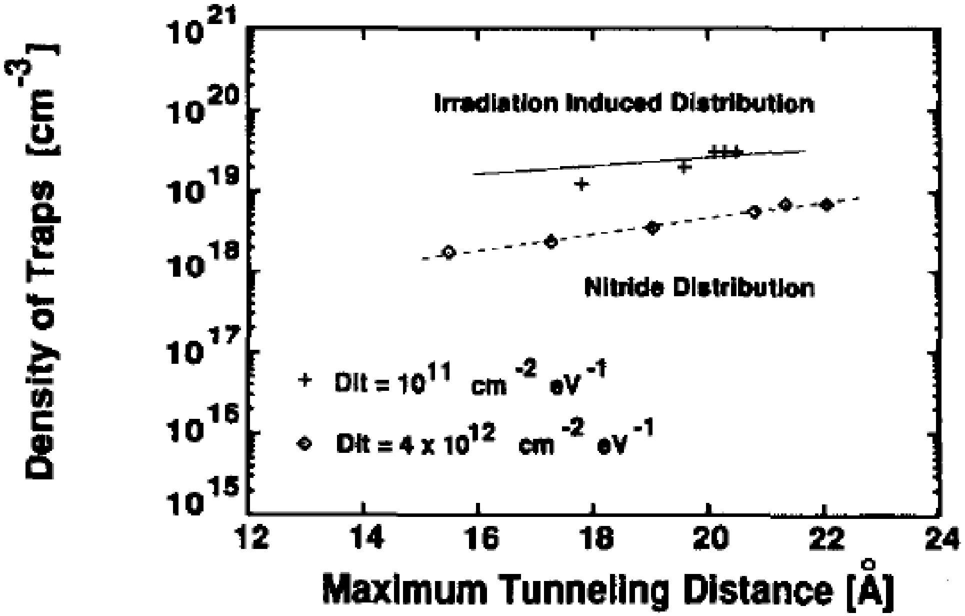


Tsuchiaki et al., IEEE TED, p. 1768, 1993
Chim et al., J. Appl. Phys., vol. 81, p. 1992, 1997
Furnemont et al., IEEE EDL, p. 276, 2007

Lateral profiles of Dit and Dot are extracted from change in transition edges of base level curves

Vertical profiling





$$\tau_T(E,x) = \frac{m_1^{*2} x (1 + \frac{1}{\alpha_1 x})}{\pi^2 \alpha_2 \hbar^3 \overline{D}_{it}} e^{\alpha_1 x} \approx \tau_0 e^{\alpha_1 x}$$

(Paulsen et al., IEEE TED, p. 1213, 1994)

Method: fill traps deeper into the oxide by increasing the time available for trapping (t_h and t_l), i.e. decreasing the frequency

Vertical profiling

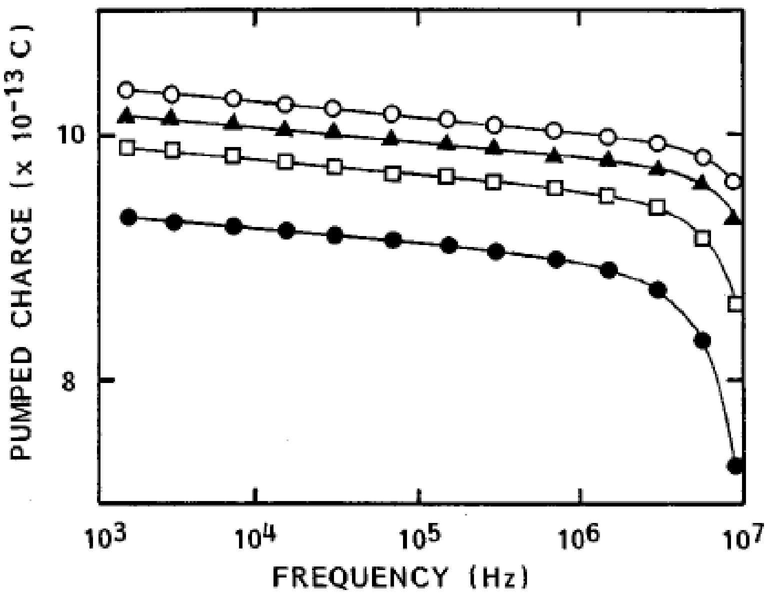


FIG. 2. $Q_{cp}(f)$ curves recorded from a stressed device under bias corresponding to the maximum of the $I_{cp}(V_l)$ at $V_{sw}=C^{ste}$ curves: ●: $V_h=0.3$ V and $V_l=-1.7$ V; □: $V_h=0.4$ V and $V_l=-2.1$ V; ▲: $V_h=0.6$ V and $V_l=-2.4$ V; ○: $V_h=0.8$ V and $V_l=-2.7$ V.

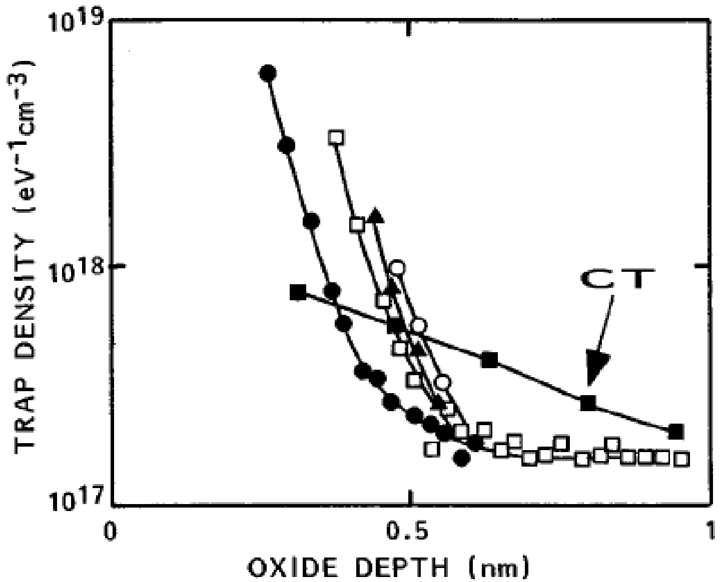


FIG. 3. Slow trap concentration profiles extracted from the data points of Fig. 2, using Eqs. (12) and (15). CP results are compared with a profile obtained using drain CT measurements.

$$Q_{cpt}=qA\Delta E\int_0^{d_{ox}}N_t(x)\Delta F(x)dx$$

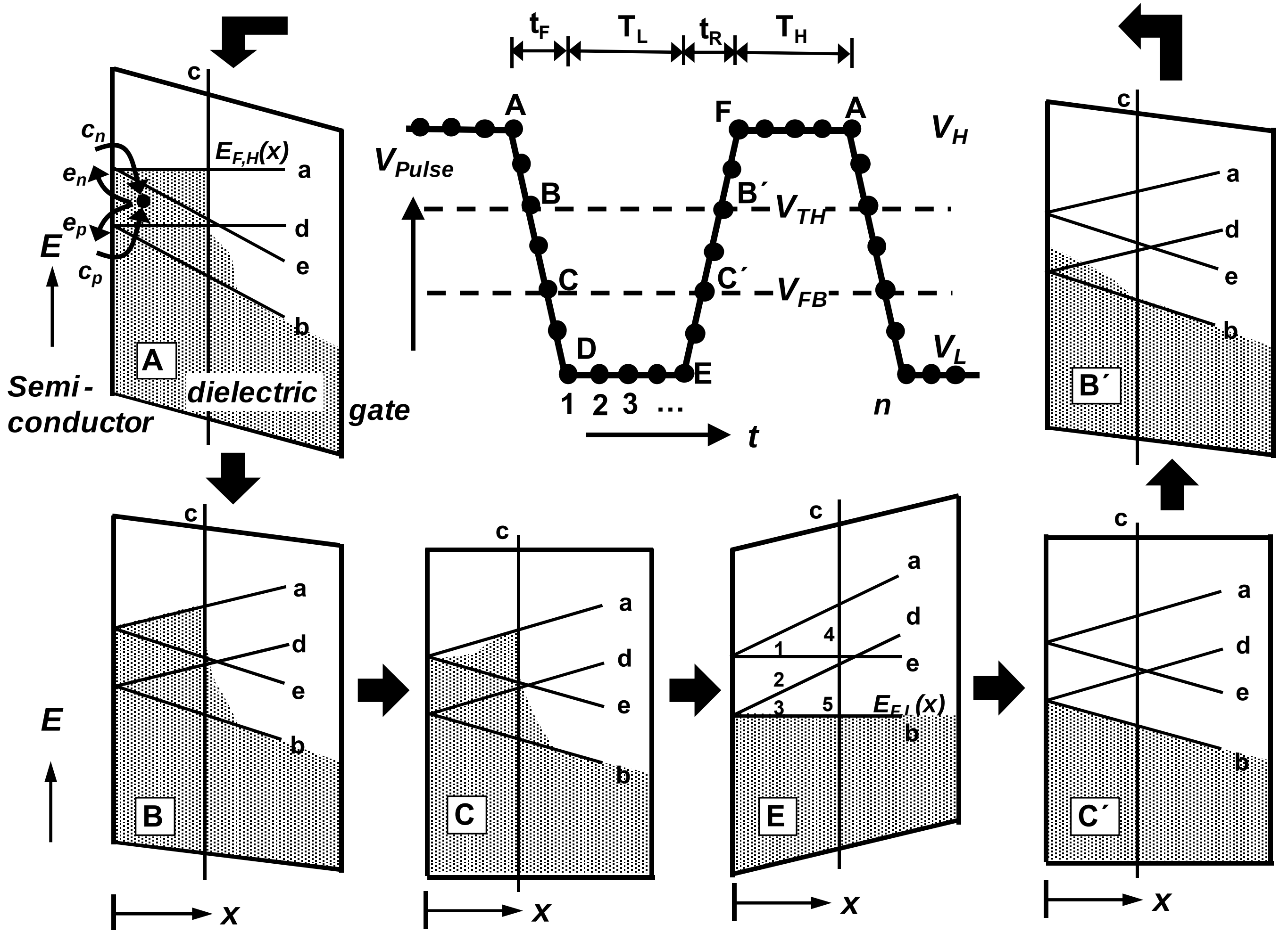
$$\sigma_n(x)=\sigma_n(0)\cdot\exp(-x/\lambda_e)$$

$$\sigma_p(x)=\sigma_p(0)\cdot\exp(-x/\lambda_h)$$

$$\Delta F(x)\simeq\frac{\{1-\exp[-c_n(x)/2f]\}\{1-\exp[-c_p(x)/2f]\}}{1-\exp\{-[c_n(x)/2f]-[c_p(x)/2f]\}}$$

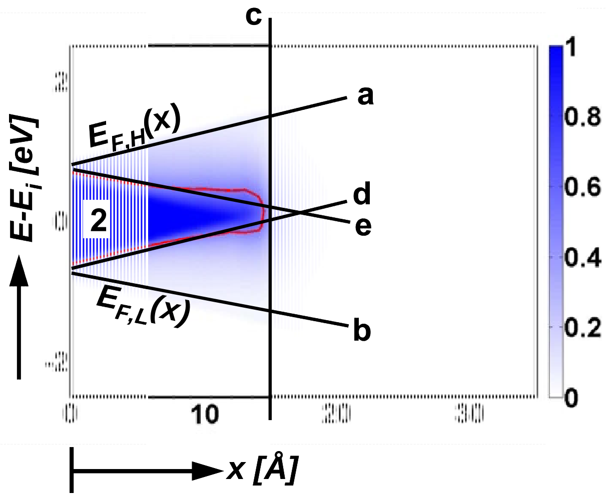
Maneglia and Bauza, J. Appl. Phys., vol 79, p. 4187, 1996

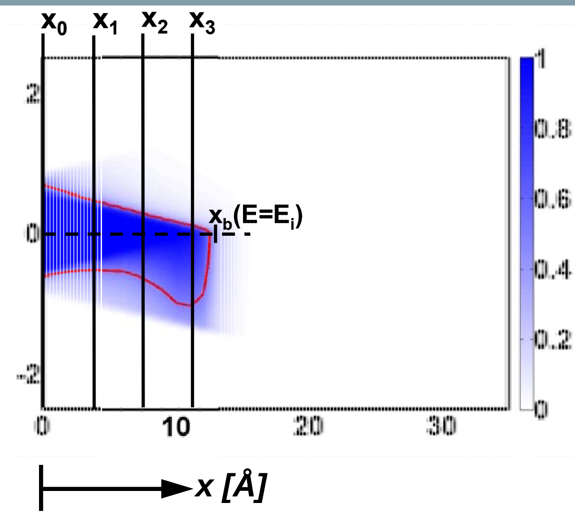
More accurate analysis

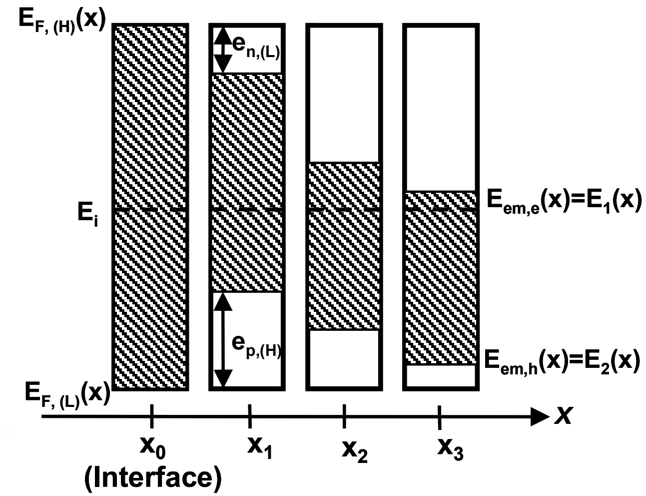


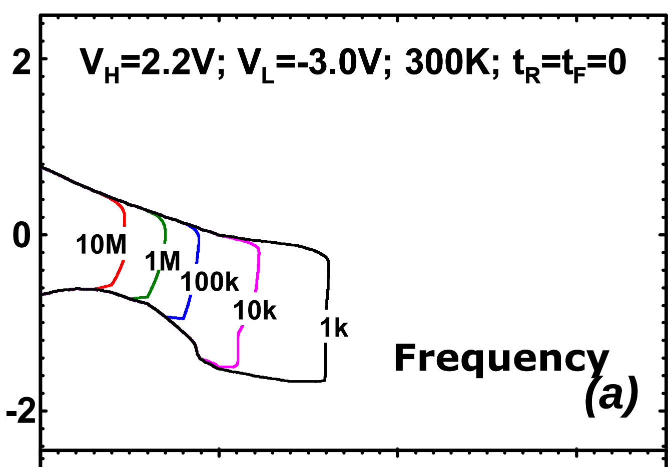
M. Masuduzzaman et al, TED Dec 2008

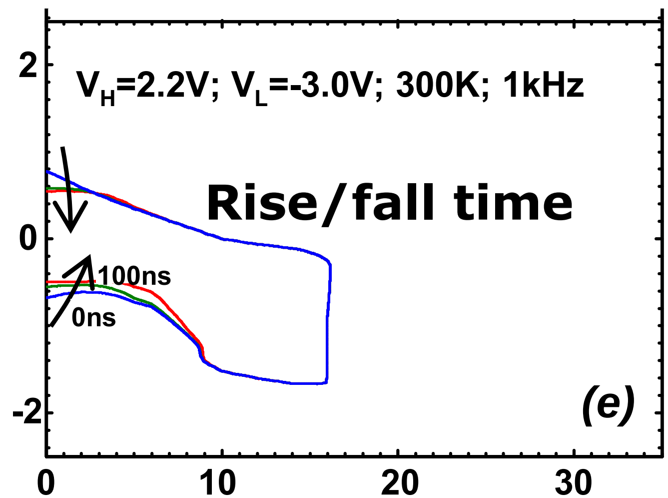
More accurate analysis











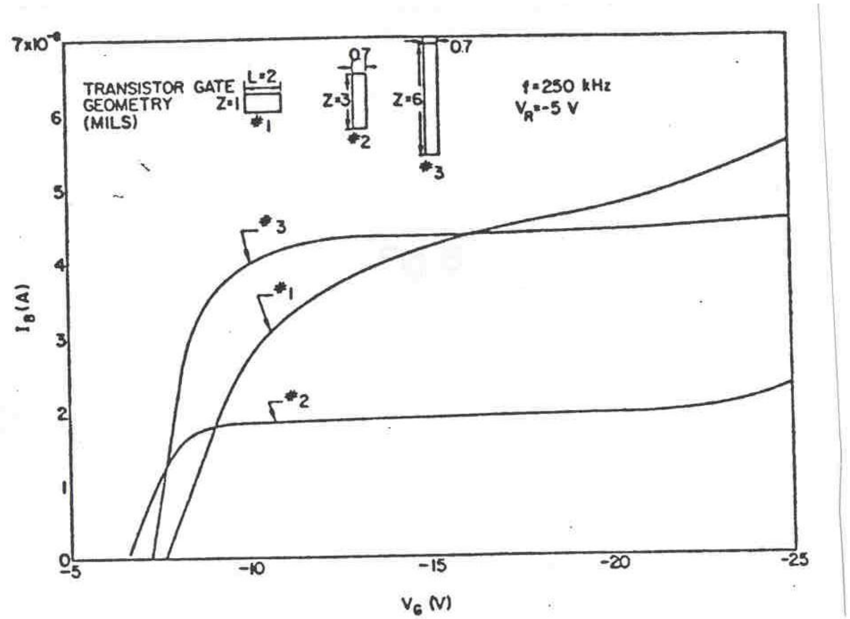
M. Masuduzzaman et al, TED Dec 2008

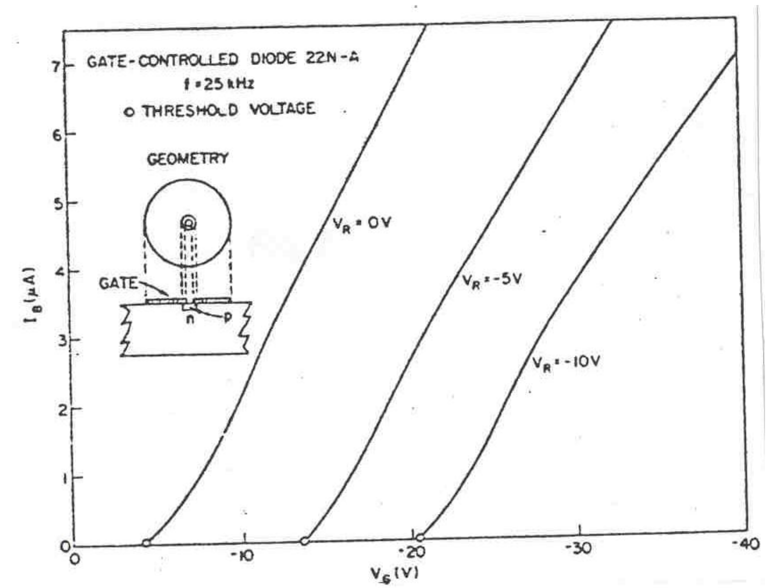
Outline

1. Introduction
2. Basic CP-principle
3. Second order model
4. Temperature dependence
5. Base level edges
6. MOSFET degradation
7. Energy distribution
8. Lateral and vertical profiling
9. Geometric components
10. Effects of oxide thickness scaling
11. Single trap characterization
12. CP in high k gate stacks

GEOMETRIC COMPONENTS

What is a geometric component ?



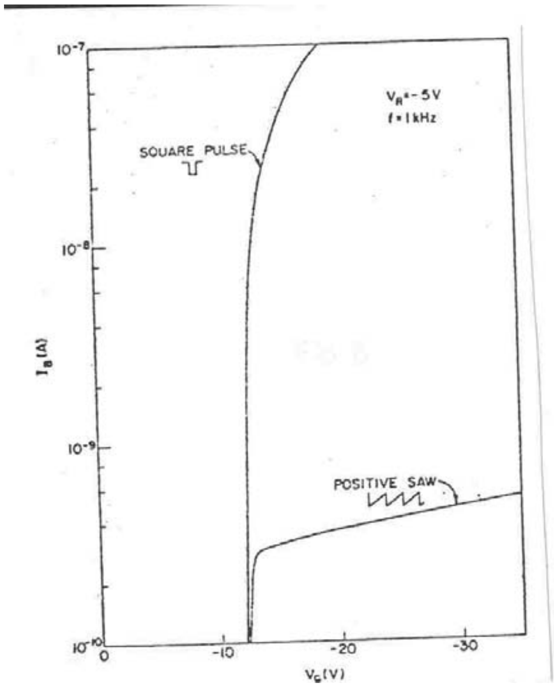


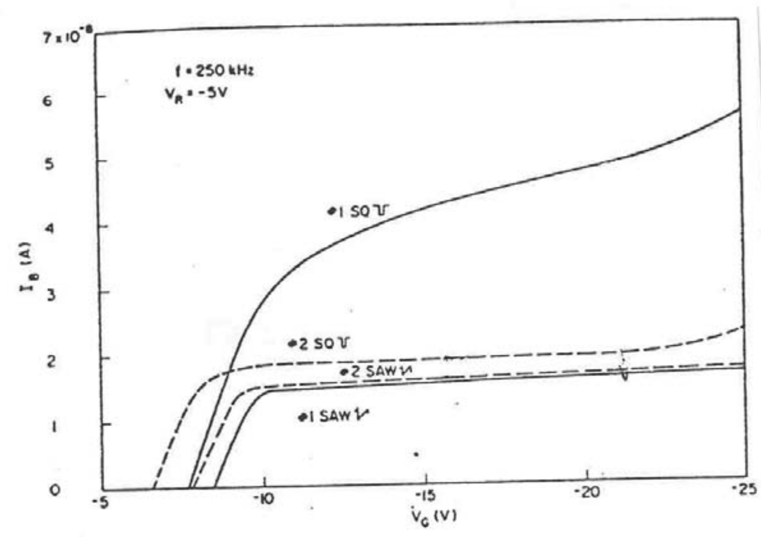
(Brugler & Jespers, 1969)

If part of the free minority carriers are recombining with majority carriers, they will be measured as Icp, and cannot be distinguished from carriers that recombine at interface traps

GEOMETRIC COMPONENTS

How to avoid a geometric component?
geometric component



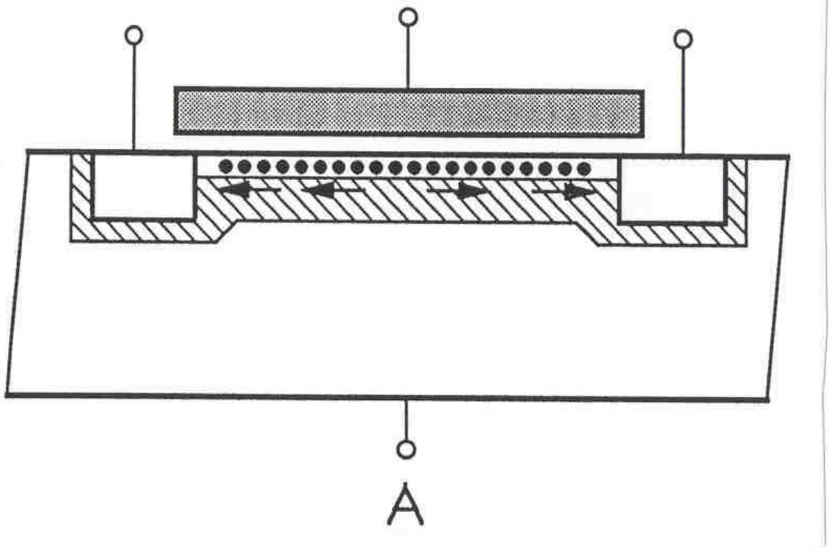


(Brugler & Jespers, 1969)

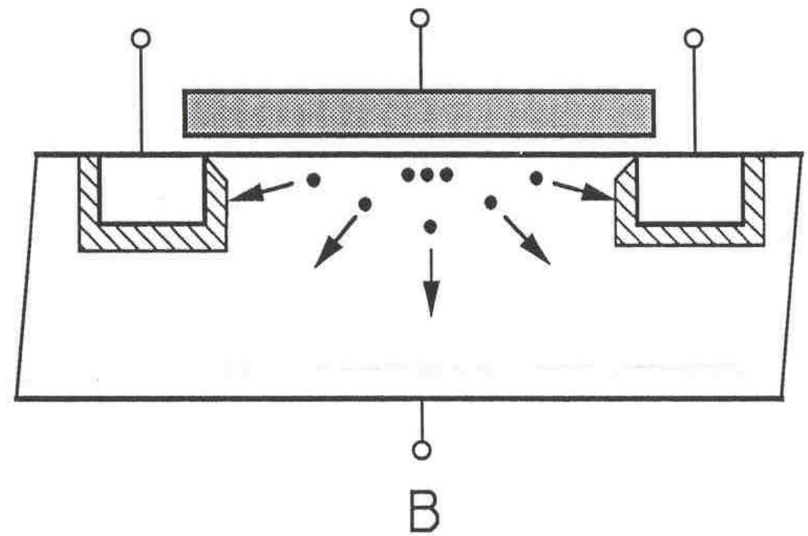
- **Use of long fall and rise times: > 10ns (e.g. triangular pulses)**
- **Use of reverse voltage at source and drain**
- **Avoid unfavorable transistor geometry: W/L>1, small**

GEOMETRIC COMPONENTS

- Switching off of a MOSFET from inversion to accumulation



Lateral drift/diffusion phase

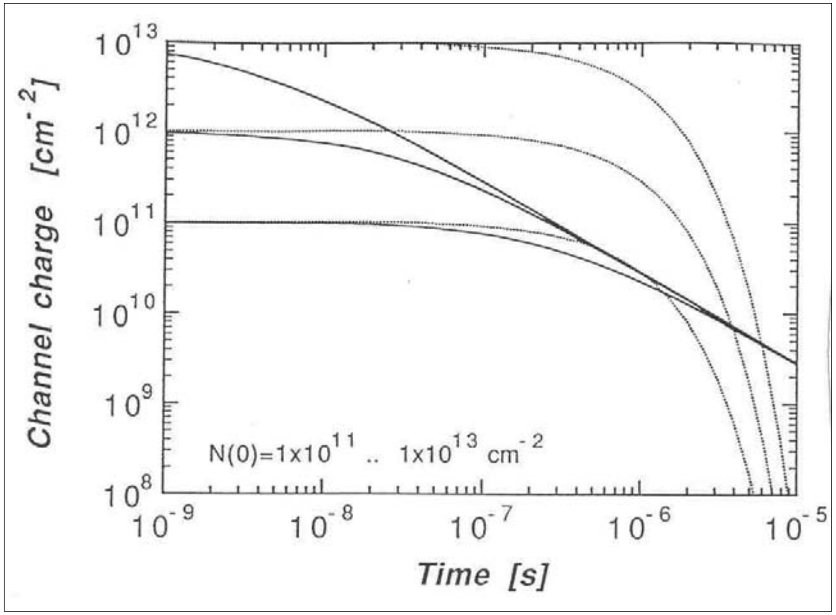


Vertical diffusion phase

G. Van den bosch et al., IEEE EDL, p. 107, 1993

GEOMETRIC COMPONENTS

Model lateral drift/diffusion phase



L = 100μm

a. ***self-induced drift:***
caused by surface potential gadient

$$\overline{n_{inv}}(t) = n_{inv}(0)\frac{t_o}{t+t_o}$$

with $t_o = \frac{\pi L_g^2 C_{ox}}{8q\mu_n n_{inv}(0)}$

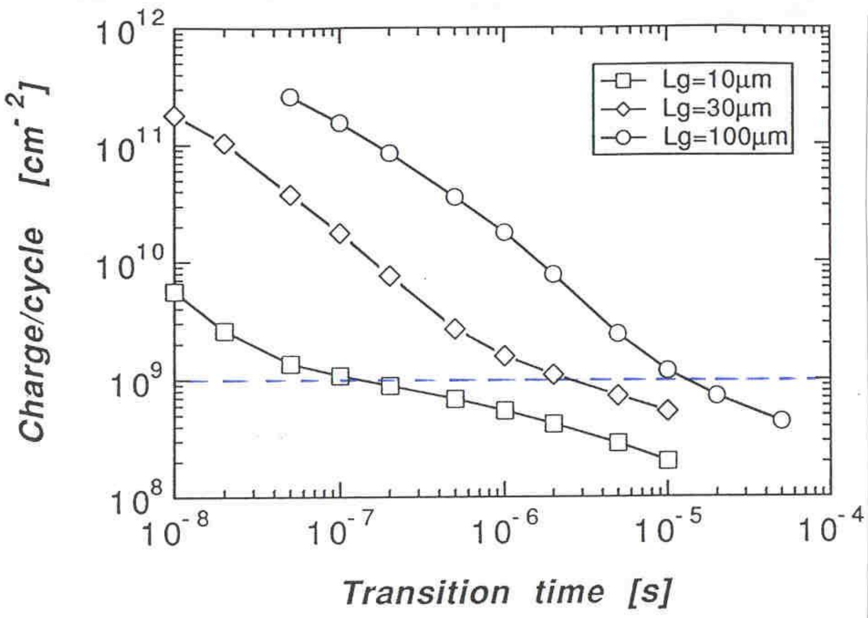
b. ***diffusion:***
caused by concentration gradient

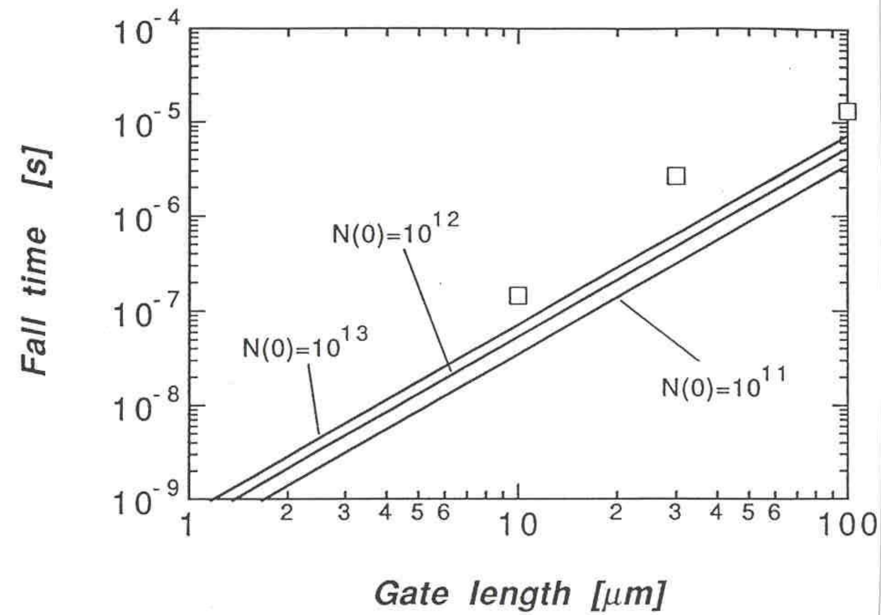
$$\overline{n_{inv}}(t) = \frac{8n_{inv}(0)}{\pi^2}\exp\left(\frac{\pi^2 D_n}{L_g^2}t\right)$$

G. Van den bosch et al., IEEE EDL, p. 107, 1993

GEOMETRIC COMPONENTS

Geometric component as a function of Channel length and pulse transition time



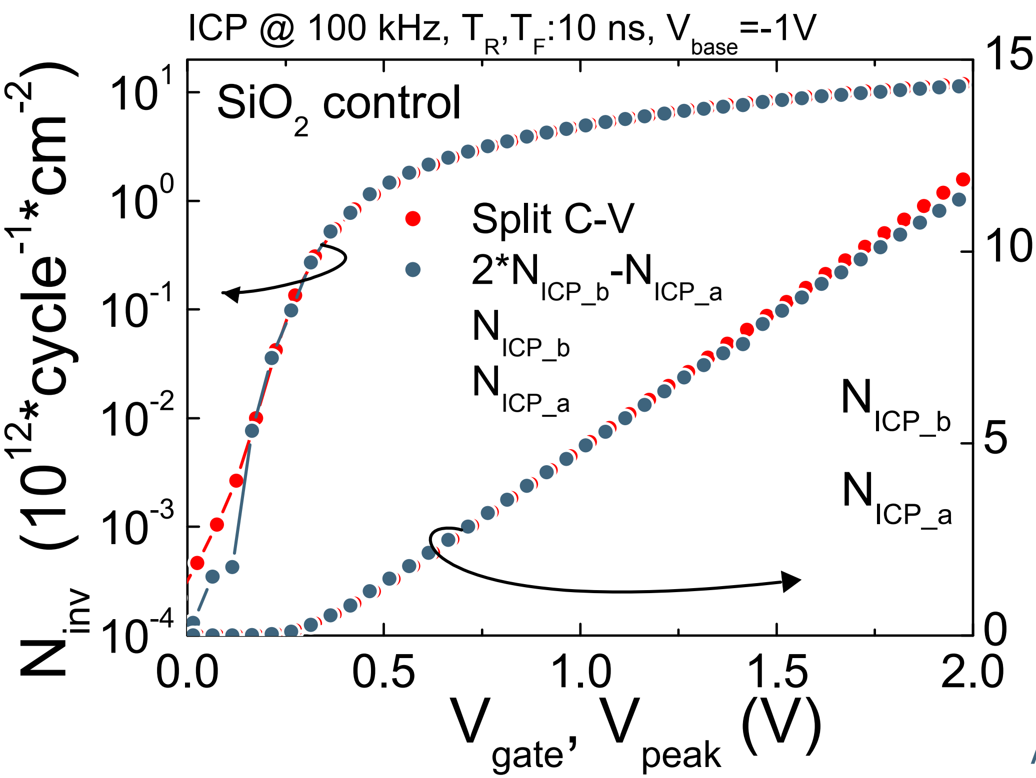


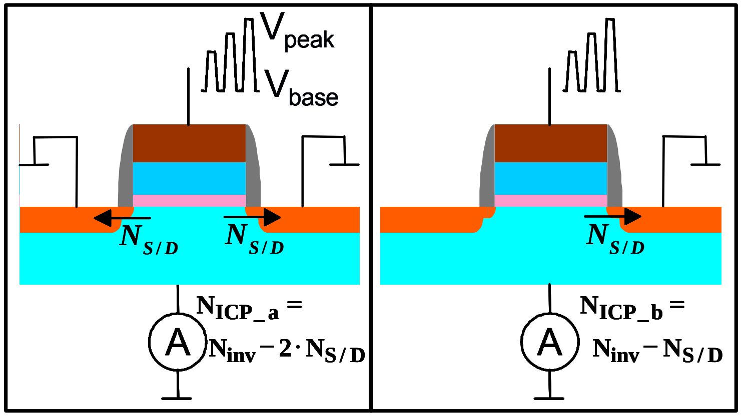
- **Influence of pulse transition time**

Influence of channel length

G. Van den bosch et al., IEEE EDL, p. 107, 1993

Inversion CP can measure inversion charge density in MOSFET





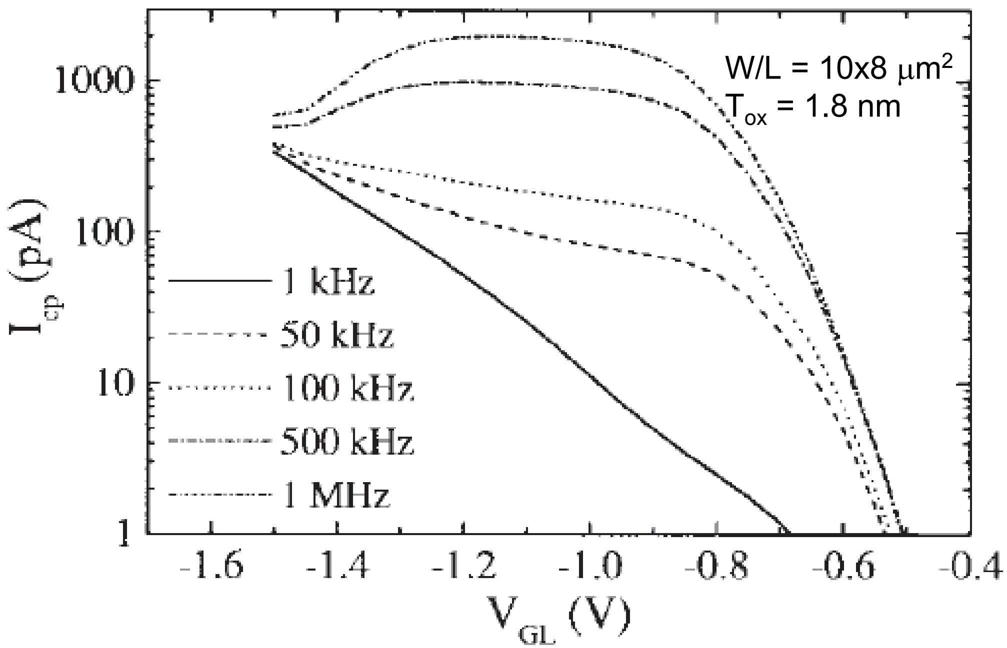
A. Kerber et al., VLSI Tech. Symp 2003

Geometric component has been used to extract inversion charge in high k MOSFET’s, as a replacement of Split-CV measurements (to avoid charge trapping effects)

Outline

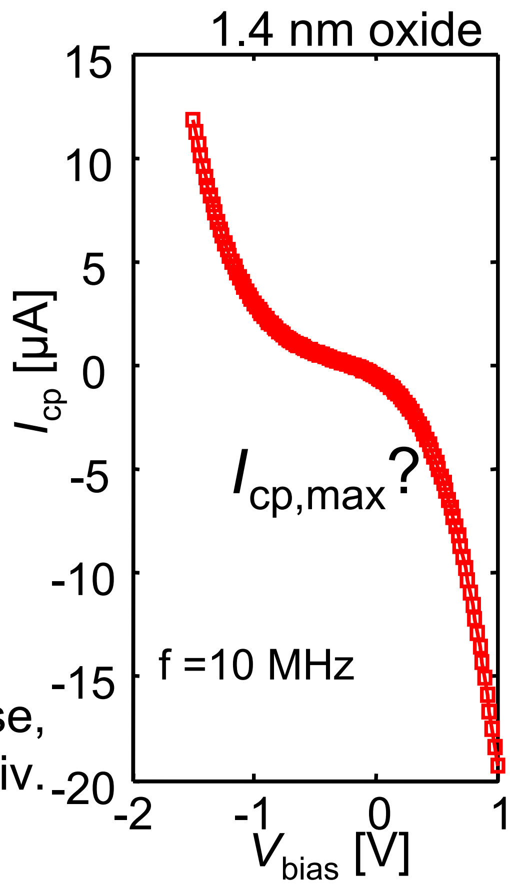
1. Introduction
2. Basic CP-principle
3. Second order model
4. Temperature dependence
5. Base level edges
6. MOSFET degradation
7. Energy distribution
8. Lateral and vertical profiling
9. Geometric components
10. Effects of oxide thickness scaling
11. Single trap characterization
12. CP in high k gate stacks

Impact of scaling oxide thickness



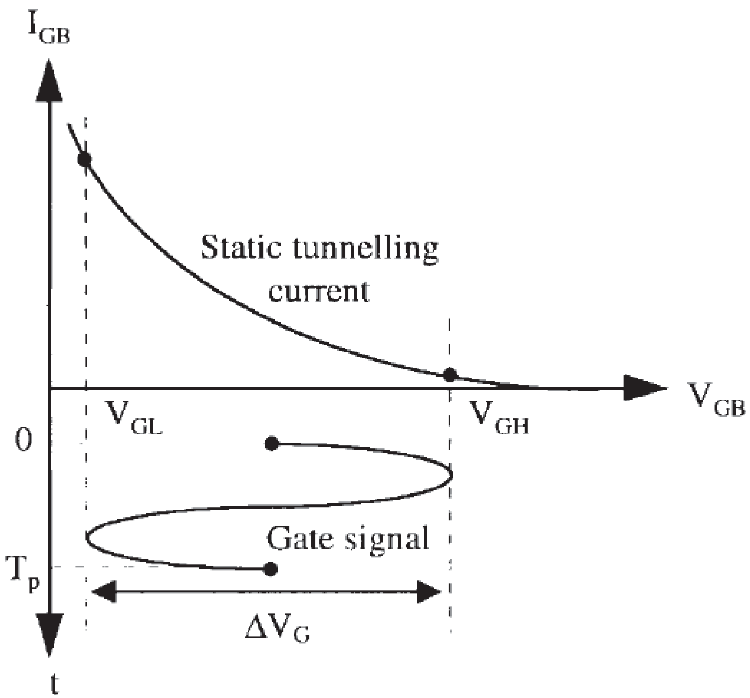
P. Masson et al.
IEEE EDL p. 92, 1999

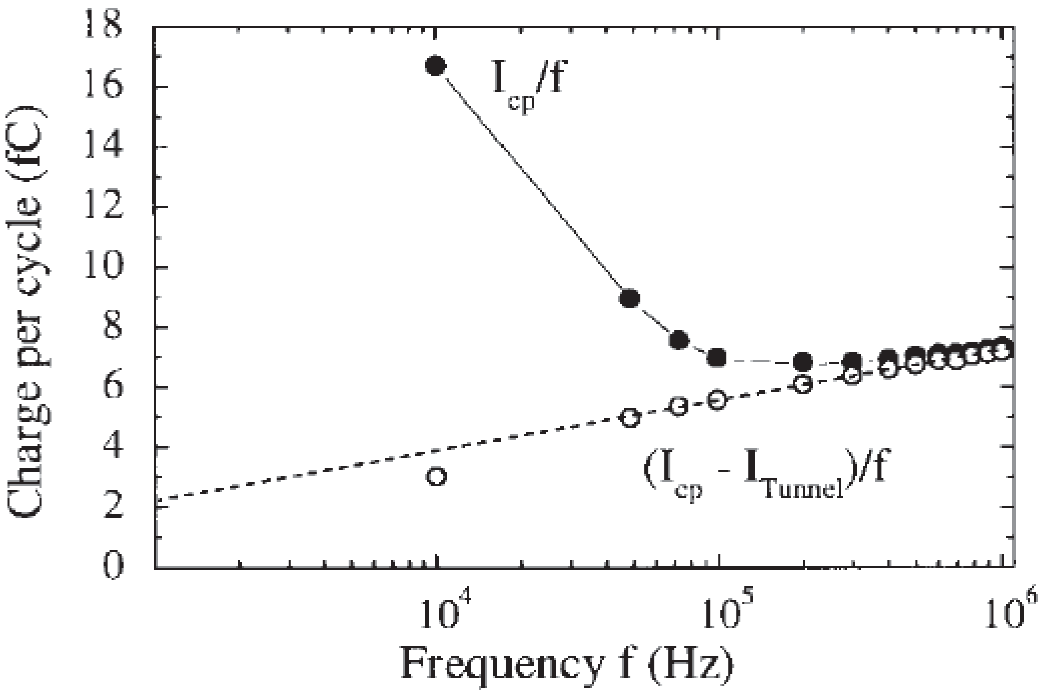
Courtesy G. Sasse,
Twente Univ.



Classical Charge Pump Technique not suitable for ultra-thin oxides due to dominance of gate leakage current !

Impact of gate leakage on CP



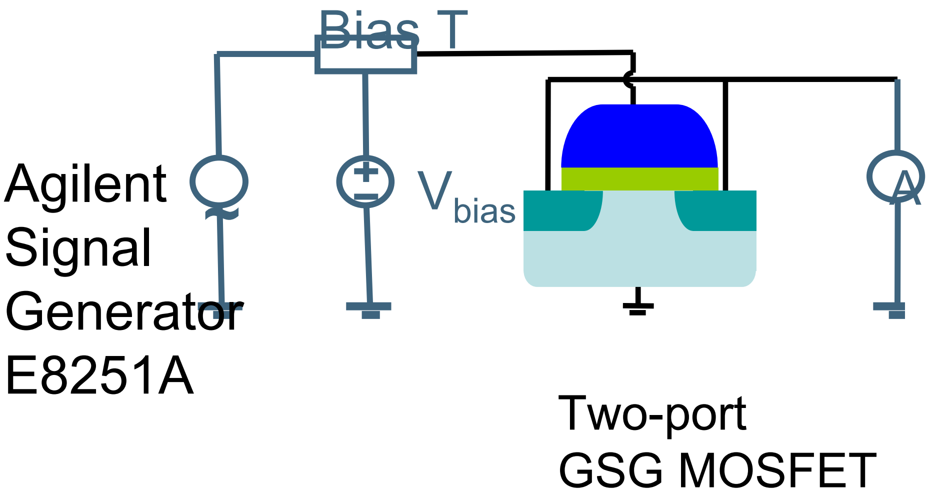


P. Masson et al., IEEE EDL p. 92, 1999

$$I_{Tunnel} = \frac{1}{T_p} \int_0^{T_p} I_{GB}\left[V_G(t)\right]dt$$

For not too high gate leakage currents, charge pumping current can be corrected for it

RF- Charge Pumping: increase the frequency



Alternative solution: increase the frequency. RF-CP

Issue with calculating the gate voltage waveform: gate impedance is dependent on voltage !
More details: G. Sasse et al. ICMTS 2005

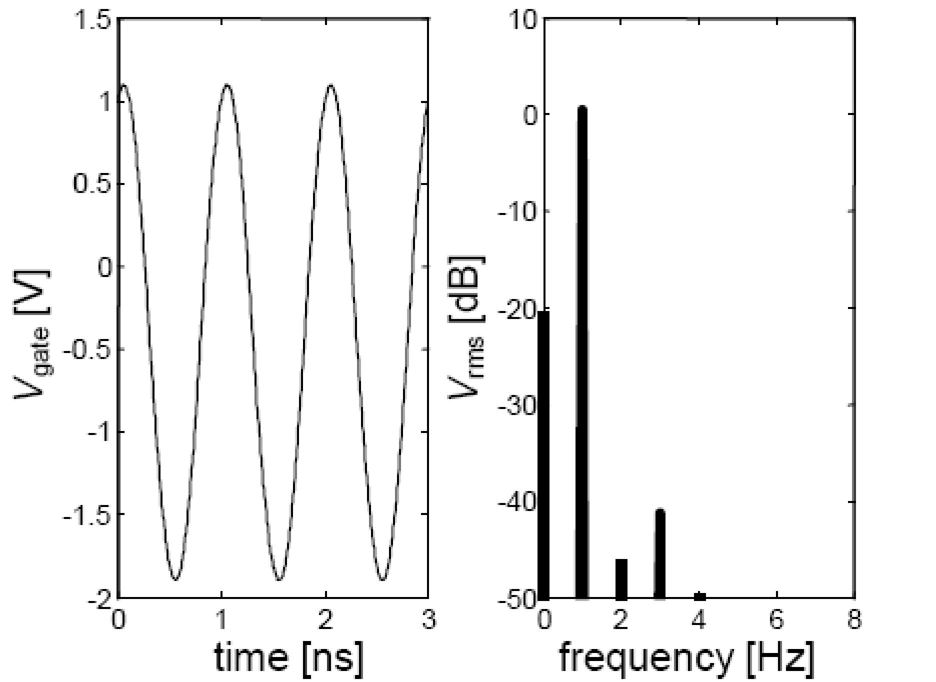
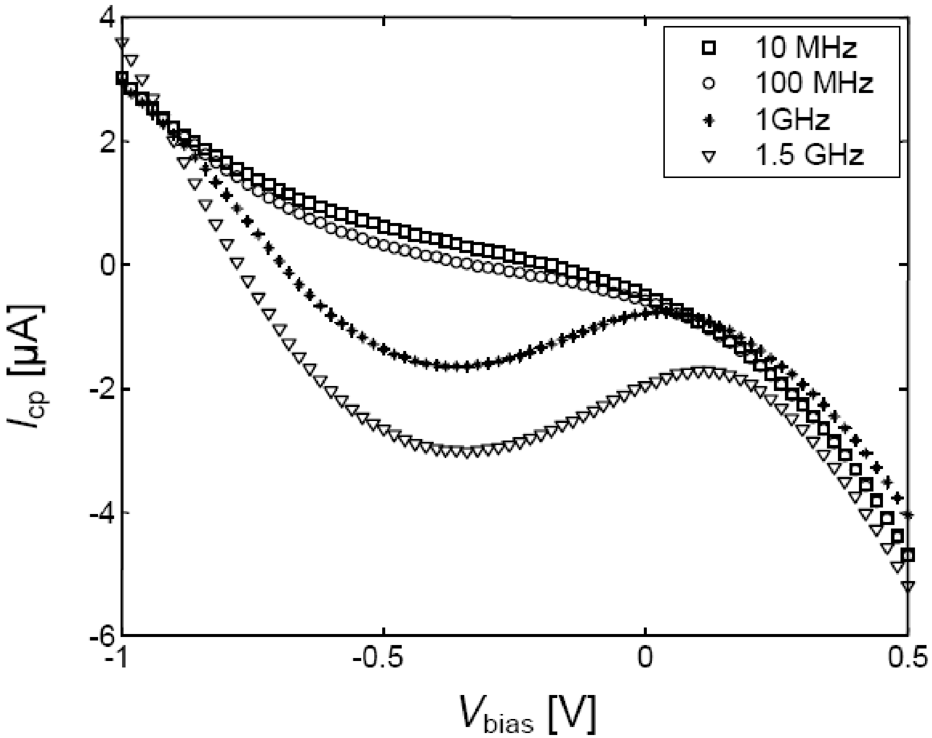
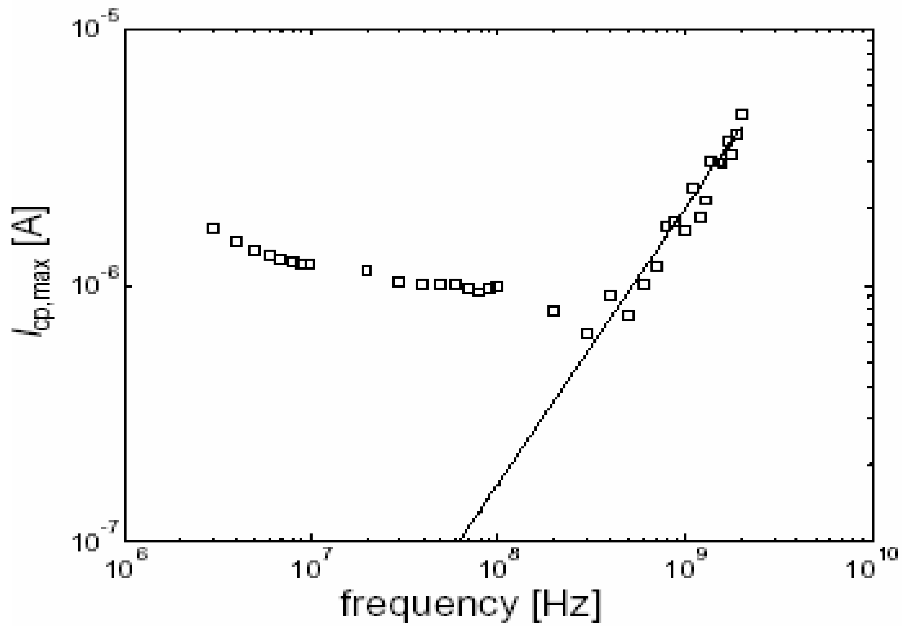


Figure 2: Estimated gate voltage waveform and the harmonic content of the time-varying signal. The voltage is calculated at a frequency of 1 GHz, input power of 9.3 dBm and applied bias voltage of -0.5 V.

RF- Charge pumping method





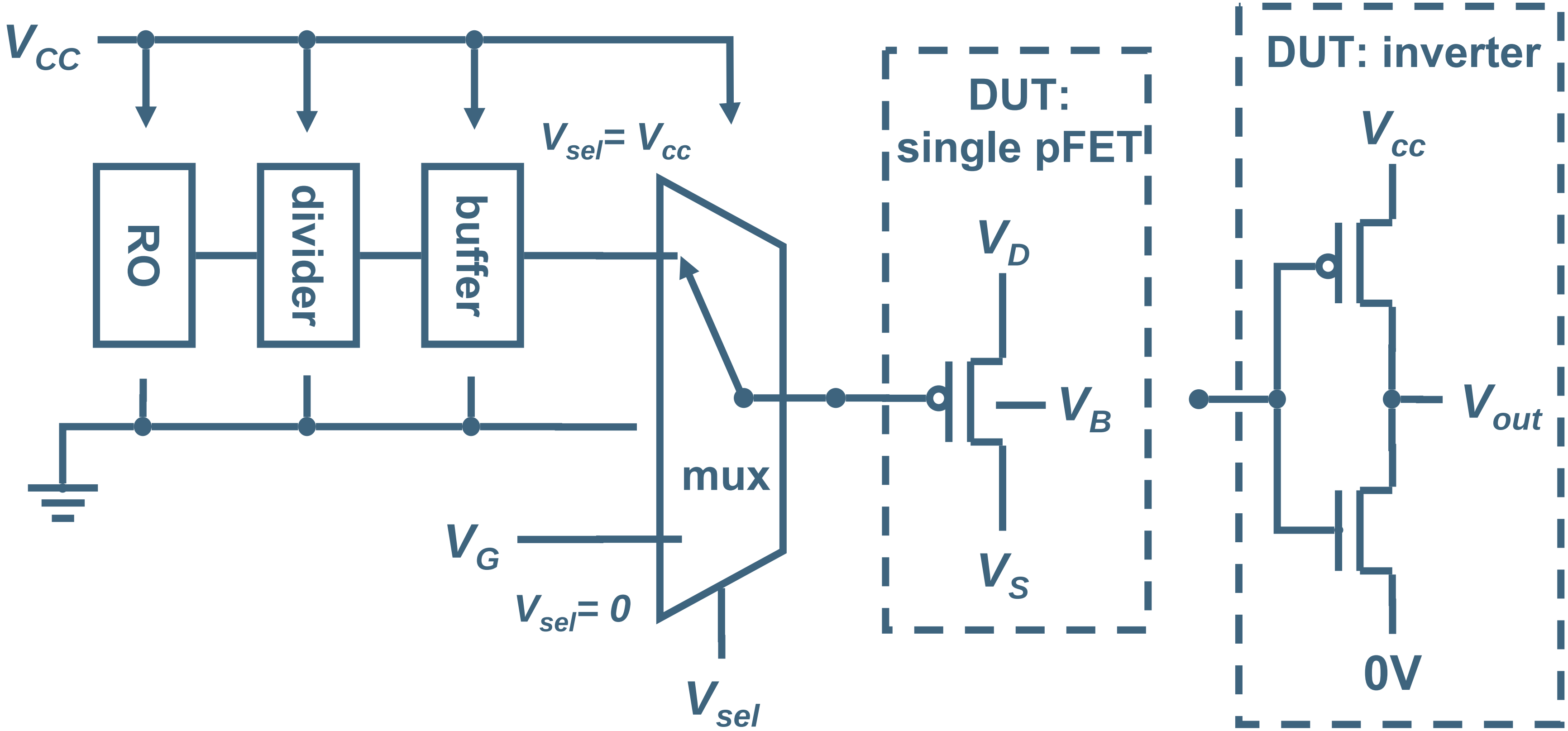
Example: tox = 1.4nm

Dit = 5×10^{10} $cm^{-2}eV^{-1}$

G. Sasse et al, ICMTS 2005

At the highest frequencies (>100MHz) the normal voltage and frequency dependence for CP is restored

On-chip circuit to study AC NBTI up to GHz range

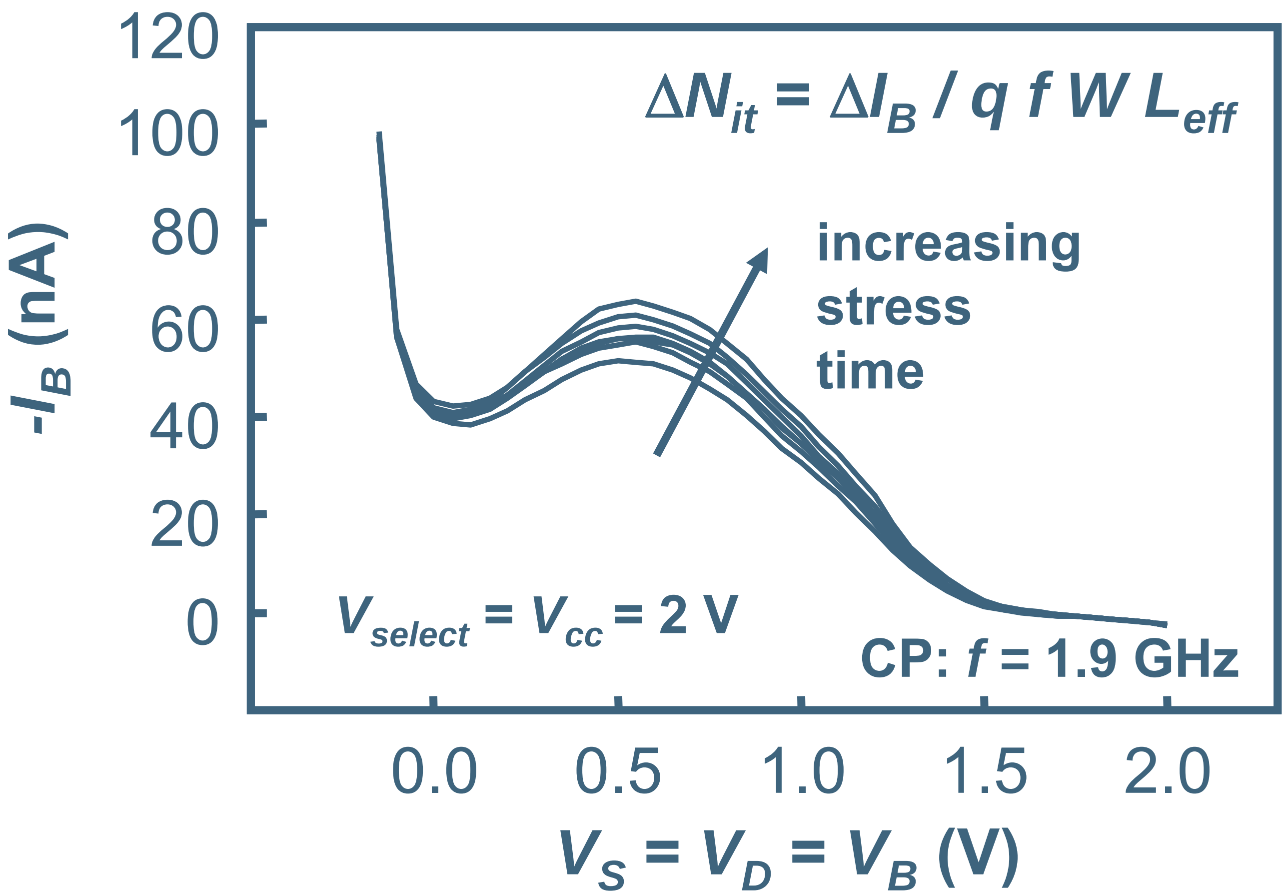


Ring Oscillator @ 2 GHz

divider = 1, 2, 4, 8, 16, 64, 256 (~7 MHz – 2 GHz)

Lower frequencies supplied externally

Circuit allows Charge Pumping in GHz range

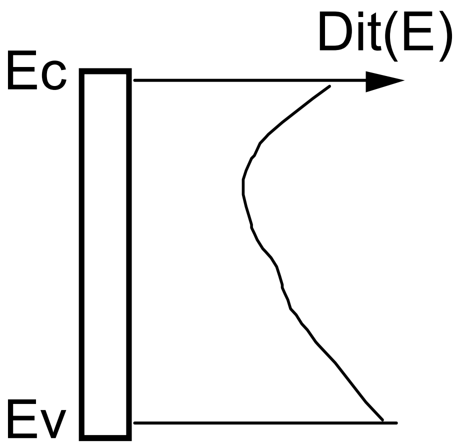


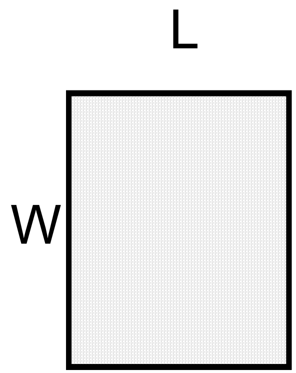
Increase in N_{it} after AC NBTI stress observed.

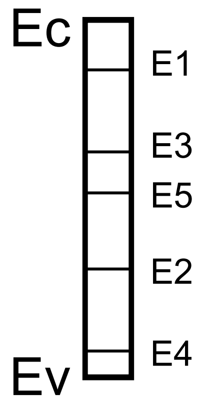
Outline

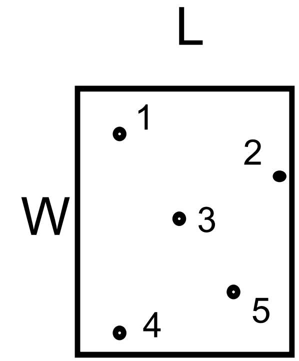
1. Introduction
2. Basic CP-principle
3. Second order model
4. Temperature dependence
5. Base level edges
6. MOSFET degradation
7. Energy distribution
8. Lateral en vertical profiling
9. Geometric components
10. Effects of oxide thickness scaling
11. Single trap characterization
12. CP in high k gate stacks

Single trap characterisation









Continuous distributions:

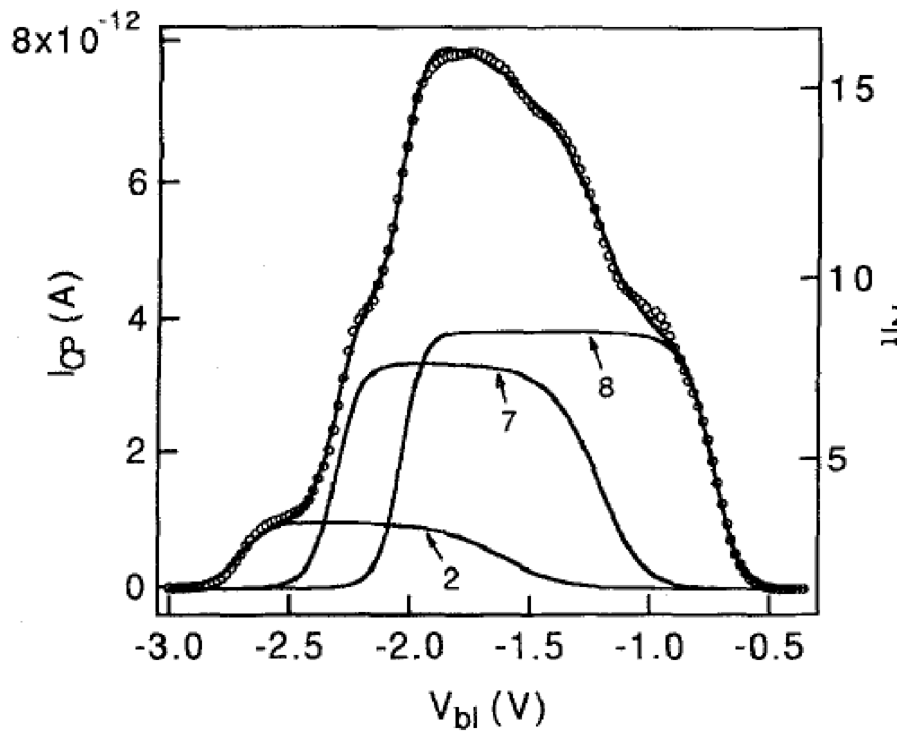
$$I_{cp} = f\,q\,A_G\,D_{it}(E)\,{}^{2}E$$

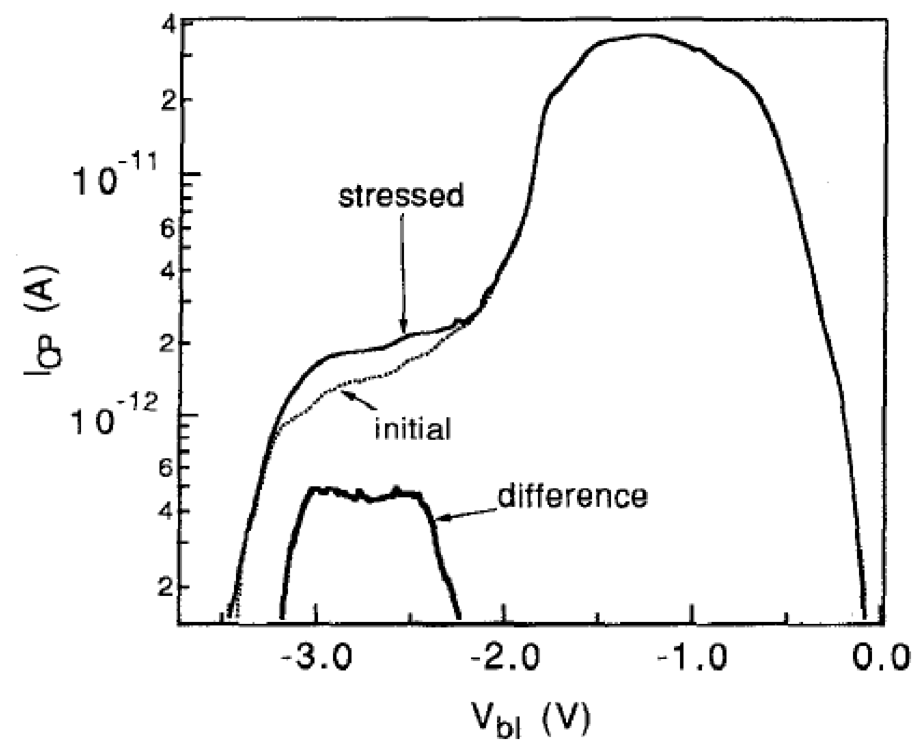
For N individual traps:

$$I_{cp} = f\,q\,N$$

$D_{it}=10^9 cm^{-2}eV^{-1}$, W=L=0.5µm
N = 2.5 traps
at 3 MHz: I_{cp} = 0.48pA/trap

Single trap characterisation



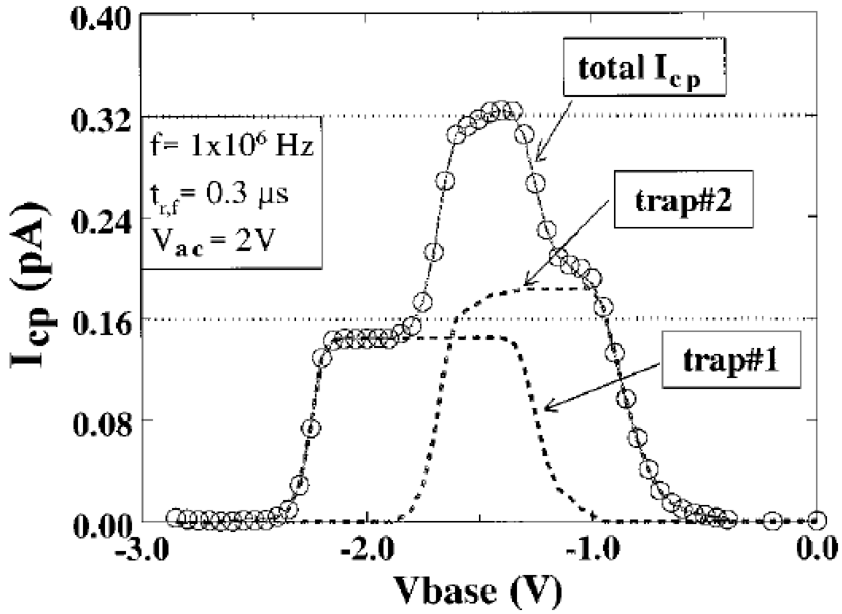


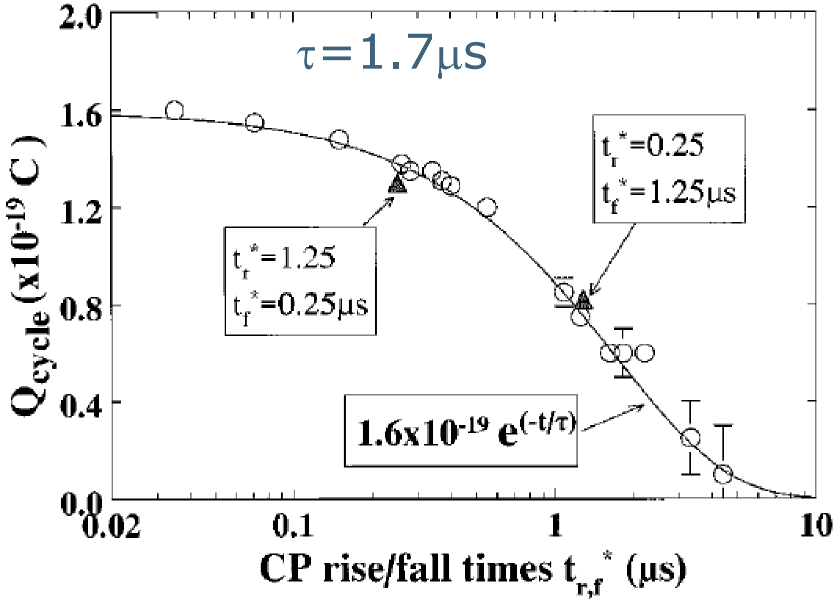
(Groeseneken et al., IEEE TED, p. 940, 1996)

Base level CP-curve, showing stepwise behavior due to individual traps

Creation of 1 single trap by a short hot carrier stress @ Vg=1.35V, Vd=3.5V

Single trap characterisation





(Saks et al, Appl. Phys. Lett., p. 1383, 1996)

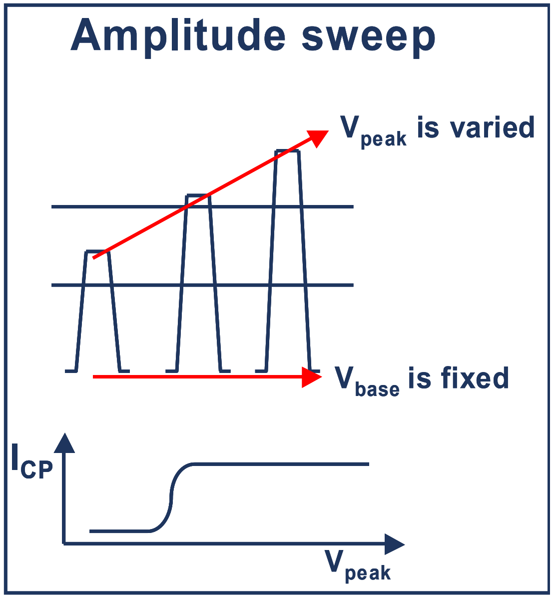
CP-curve for a MOSFET with 2 traps only (f=1MHz)

Qcp vs. rise/fall time shows exponential decay by SRH-electron emission with τ=1.7μs

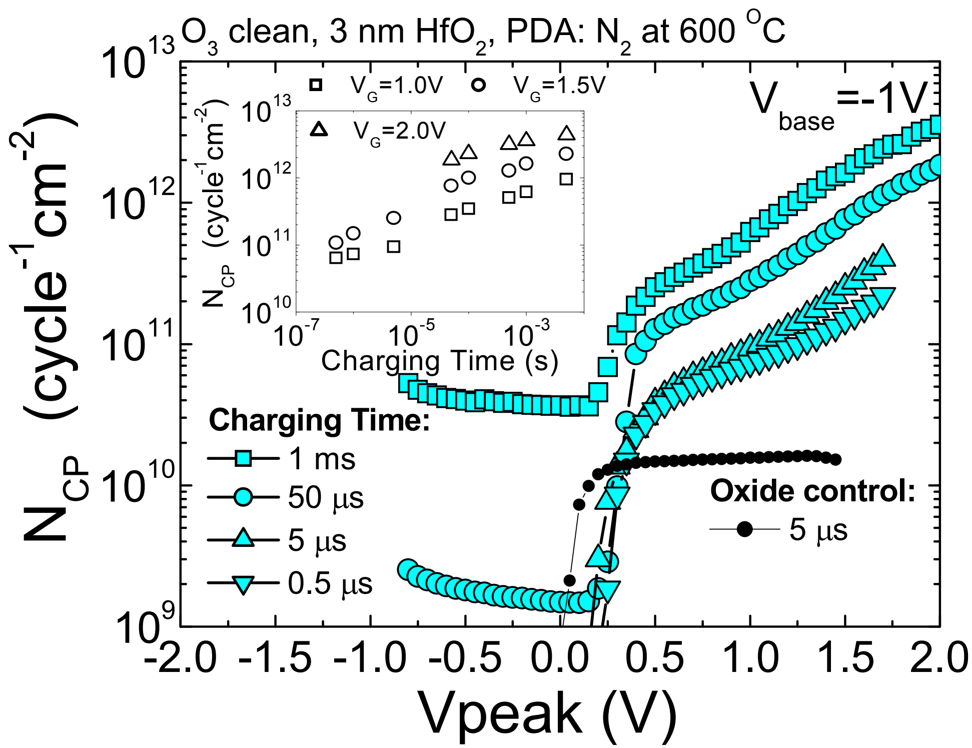
Outline

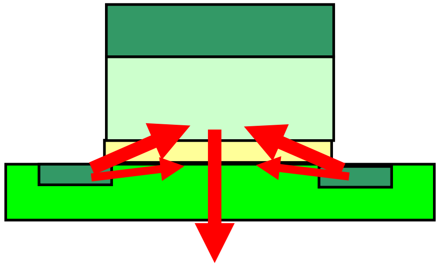
1. Introduction
2. Basic CP-principle
3. Second order model
4. Temperature dependence
5. Base level edges
6. MOSFET degradation
7. Energy distribution
8. Lateral and vertical profiling
9. Geometric components
10. Effects of oxide thickness scaling
11. Single trap characterization
12. CP in high k gate stacks

Charge pumping on high k



Amplitude sweep (at fixed base level) is used to measure charge in HfO_2 layer





A. Kerber, INFOS2003

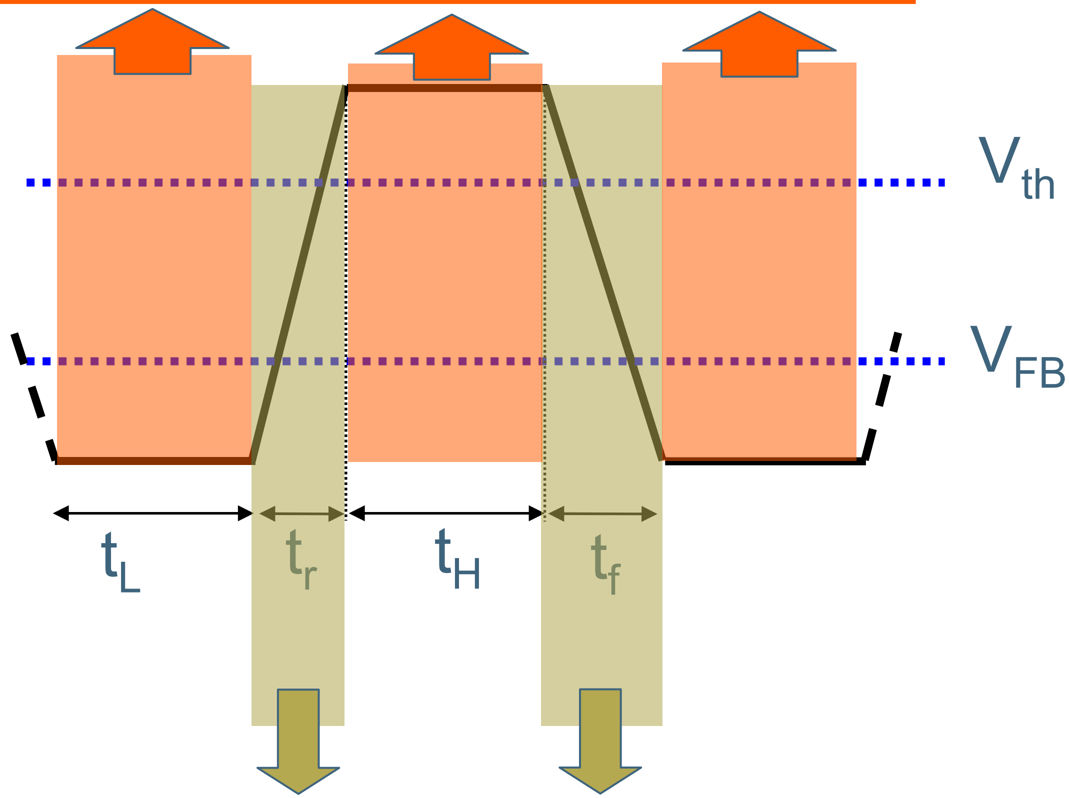
Amplitude sweep CP-level is not constant for toplevel > Vt: points to bulk trap states contributing to CP-signal !

Amount of traps measurable depends on amplitude and frequency of the CP gate voltage pulse

N_{CP} ~ $3\cdot10^{12}$ cm^{-2} for a 1 ms charging time and a V_{peak} of 2V)

Bulk trap Charge Pumping

Bulk trap charge pumping happens during the high and low time of the pulse



Classical charge pumping happens during rise and fall time only

t_L = time at low V_G
t_H = time at high V_G
t_r = rise time
t_f = fall time

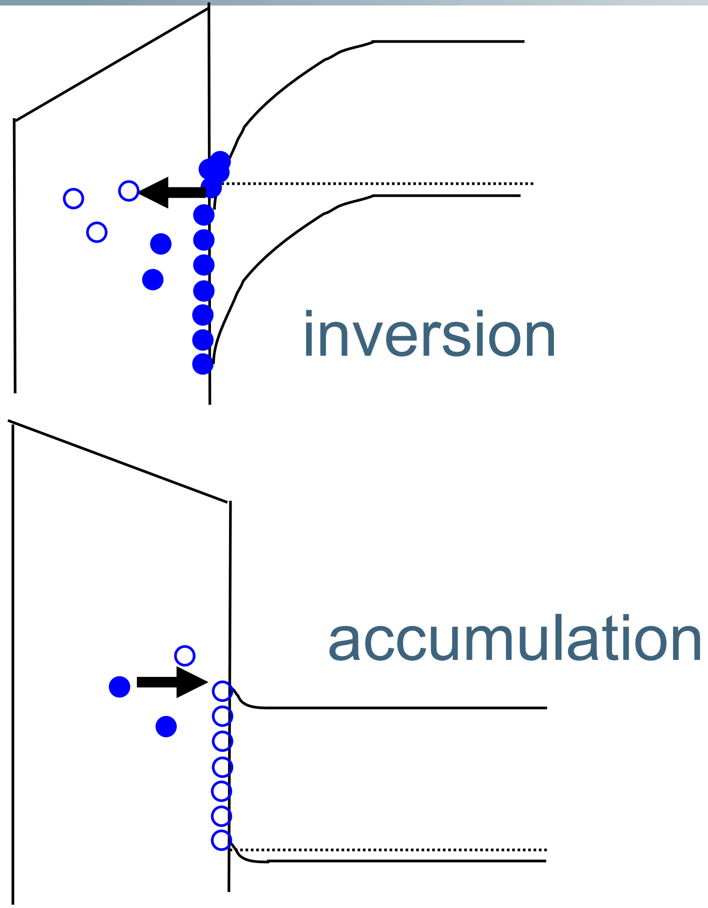
Q_c = pumped charge per cycle
= I_{CP}/f
= qAN_{it}
= **independent** of frequency in **classical** formula

but Qc is **NOT** independent of frequency when bulk states are pumped!

Charge pumping on high k

During t_H bulk states can also be occupied through tunneling

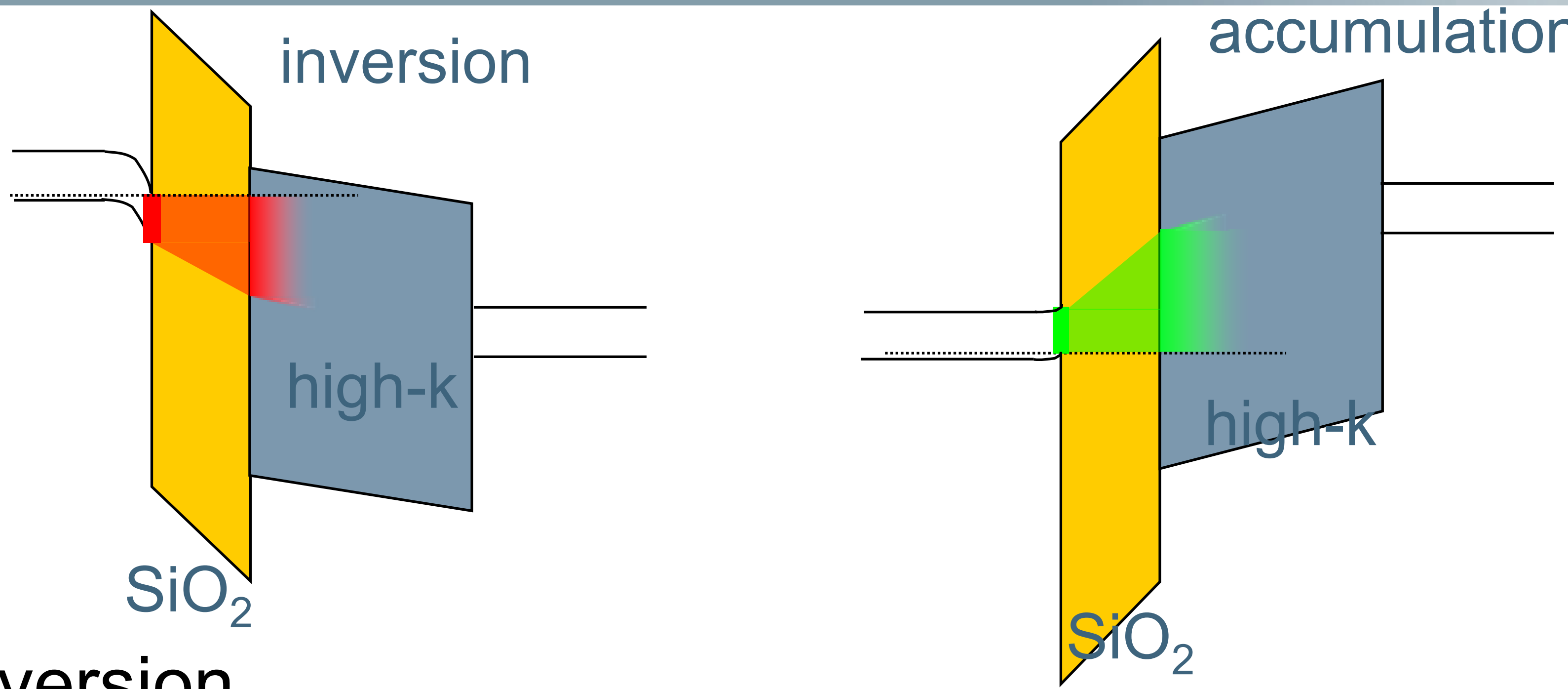
During t_L bulk states can also be emptied through tunneling and recombine



Bulk charge pumping

- happens during t_H and t_L
- The longer t_H and t_L, the further away from the interface traps can participate in the charge pumping
- charge/cycle depends t_H,t_L, in other words on frequency

Charge pumping can sense high-k trap density



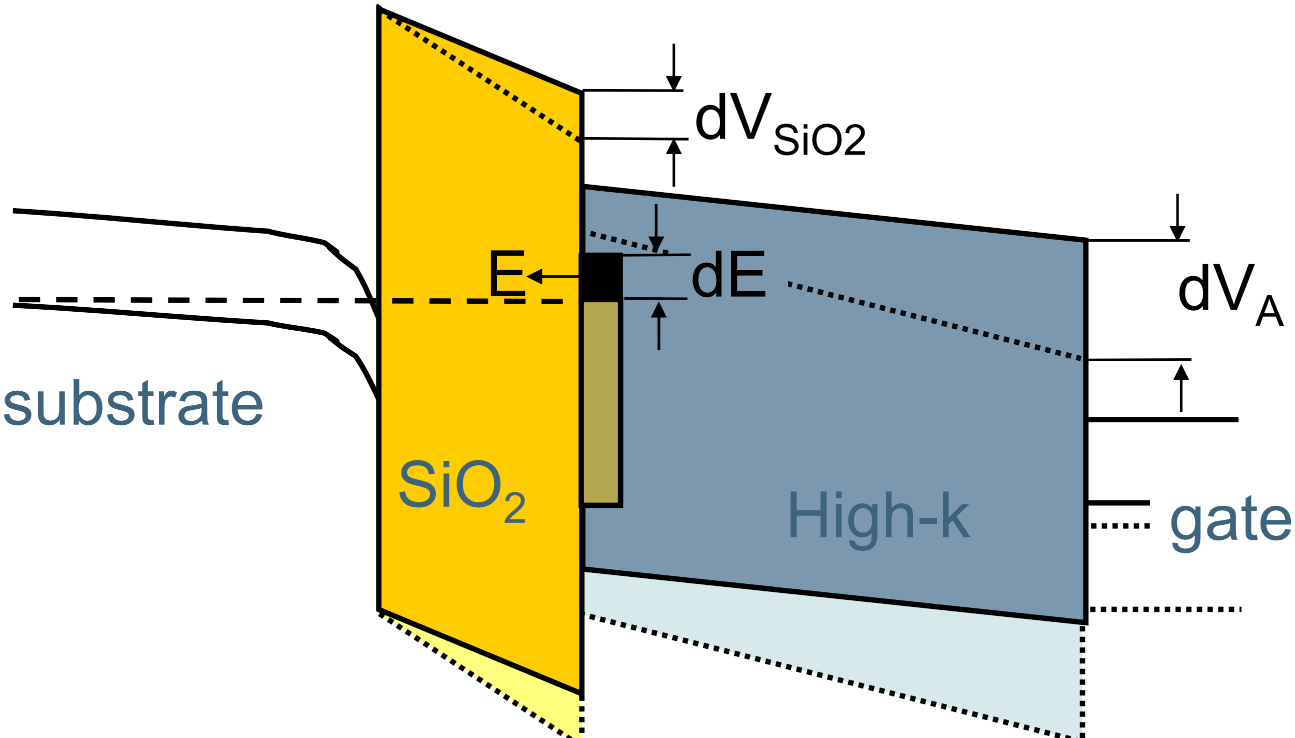
Inversion

interface traps are filled with electrons during voltage ramp (conventional CP) Pumping)
High-k bulk traps are also filled through tunneling during t_H

Accumulation

Interface states are emptied and electrons recombine during voltage ramp-down
High-k bulk states are also emptied by tunneling and electrons recombine during t_L

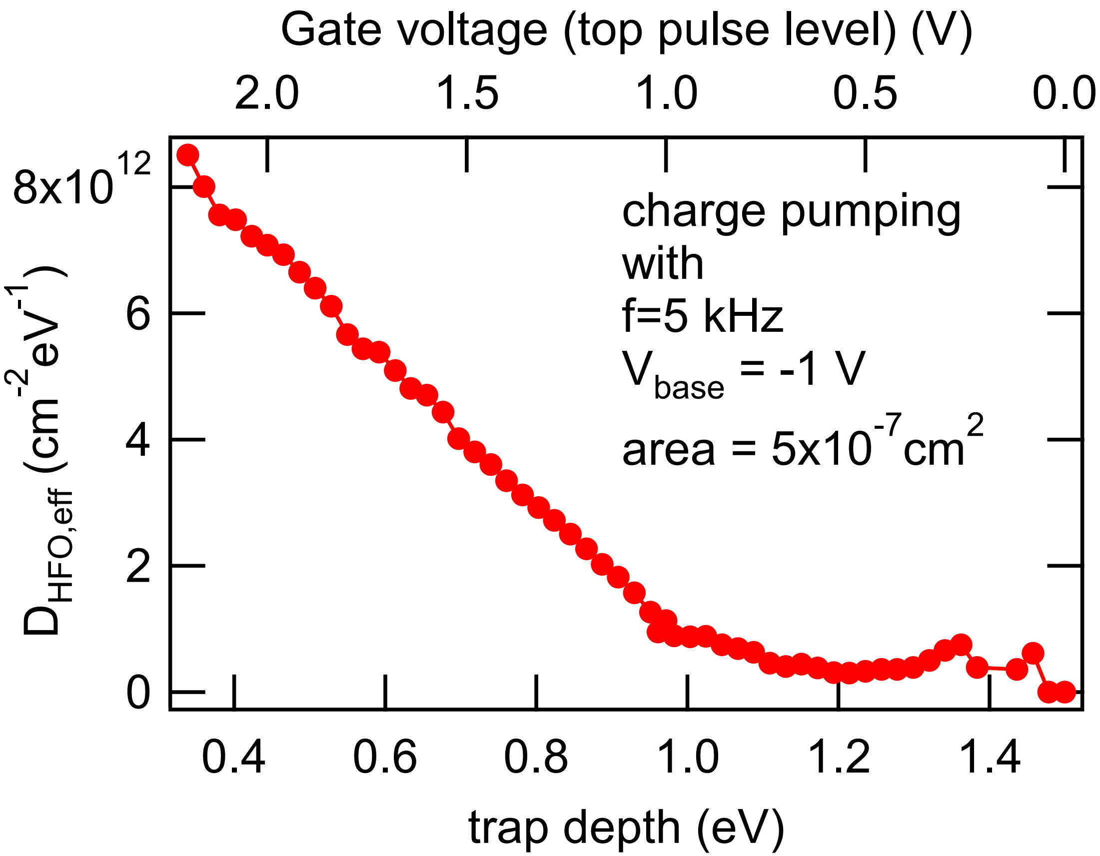
Variable amplitude Charge Pumping can scan trap energy spectrum

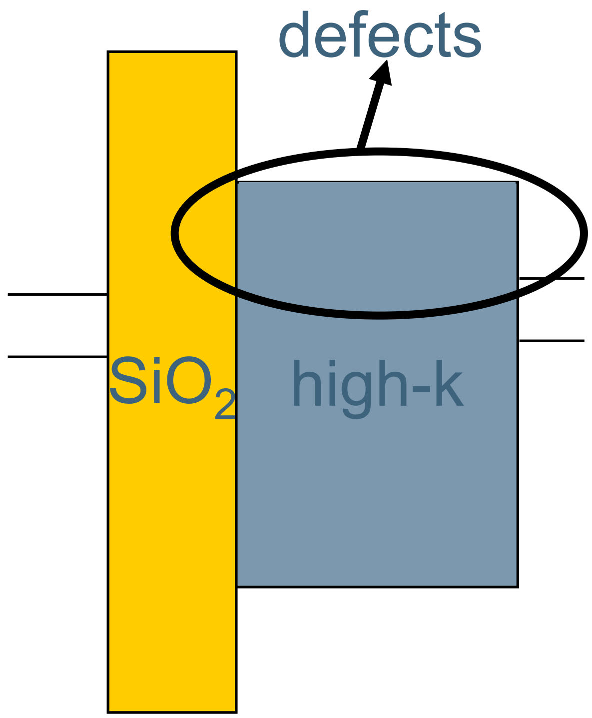


Traps in dE = extra traps measured when changing the amplitude over dV_A

Traps close to SiO_2/high-k are sensed → Measure D_{eff} (in $eV^{-1}cm^{-2}$)

Defect 'band' near HfO_2 conduction band edge





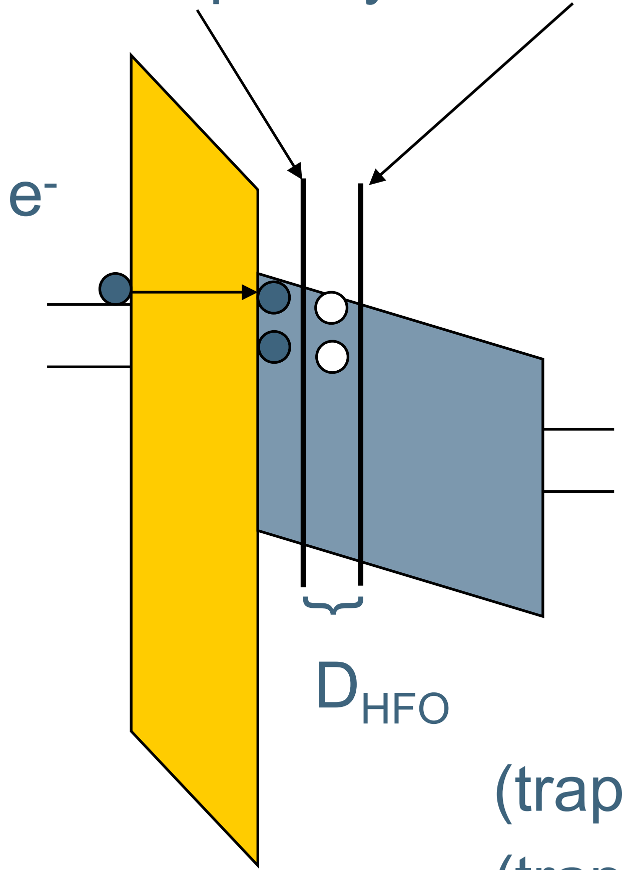
Trap depth from top of HfO_2 conduction band

Shallow traps are sensed at high gate voltage

These are responsible for V_{th}-hysteresis

Charge pumping to sense high-k bulk traps

frequency 1 > frequency 2

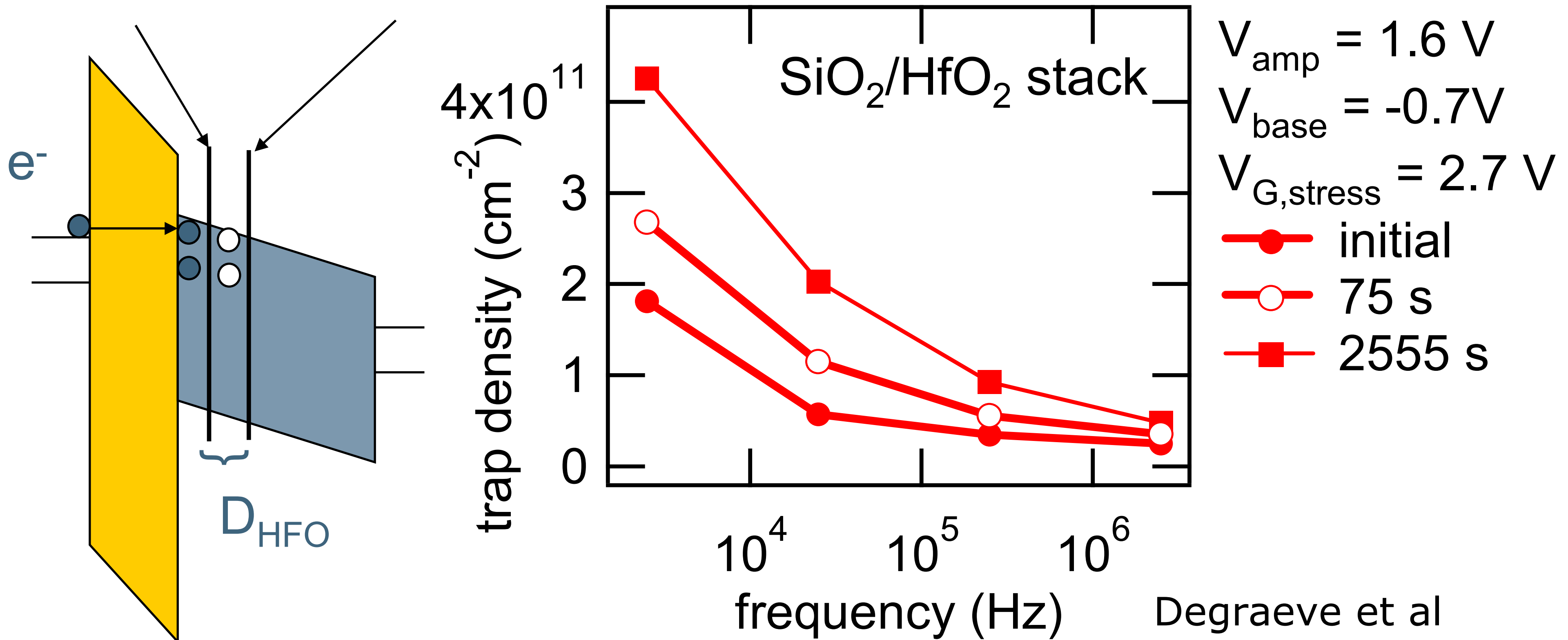


- At low frequency
  - bulk HfO_2 traps are sensed
- At very high frequency
  - no bulk traps
  - pure Si/SiO_2 interface states
  - =conventional Charge Pumping

(trap density at f_2)
– (trap density at f_1)

= fraction of bulk HfO_2 trap density

Frequency scan for bulk states

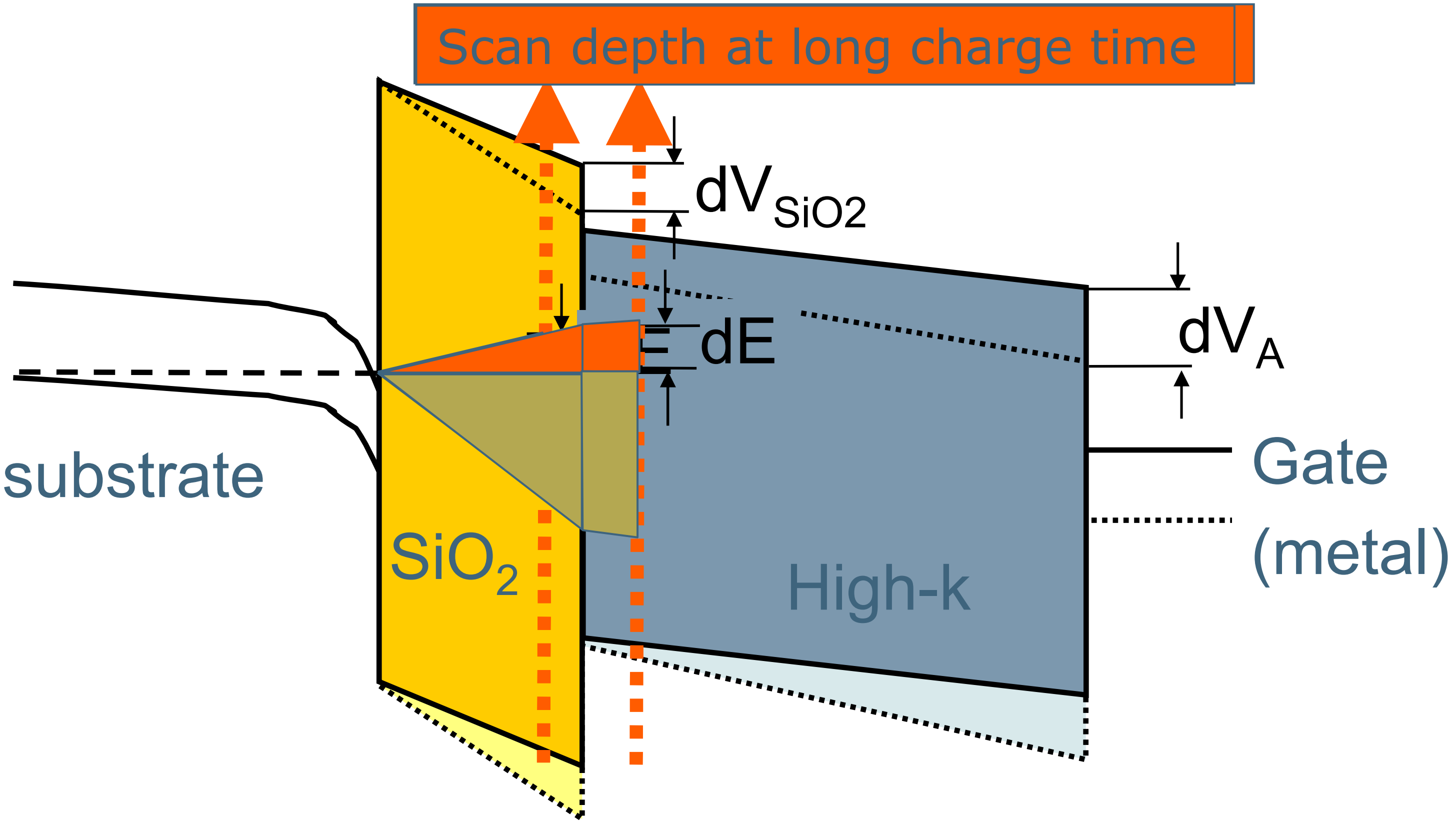


Degraeve et al IEDM 2003

Interface states at high frequency

The lower the frequency, the more bulk traps

Principle of distance and energy scanning

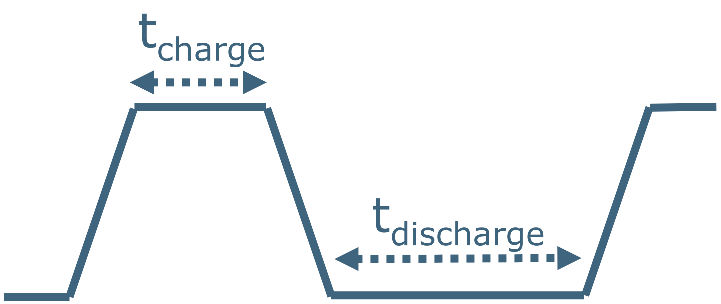


Two main parameters: **charge time and amplitude**

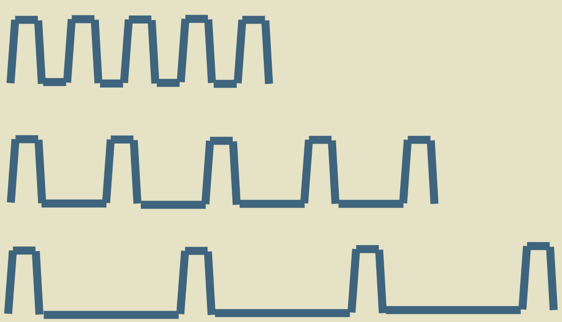
1) the trap distance from the injecting interface → t_{ch} controlled
2) the trap energy level → V_A controlled

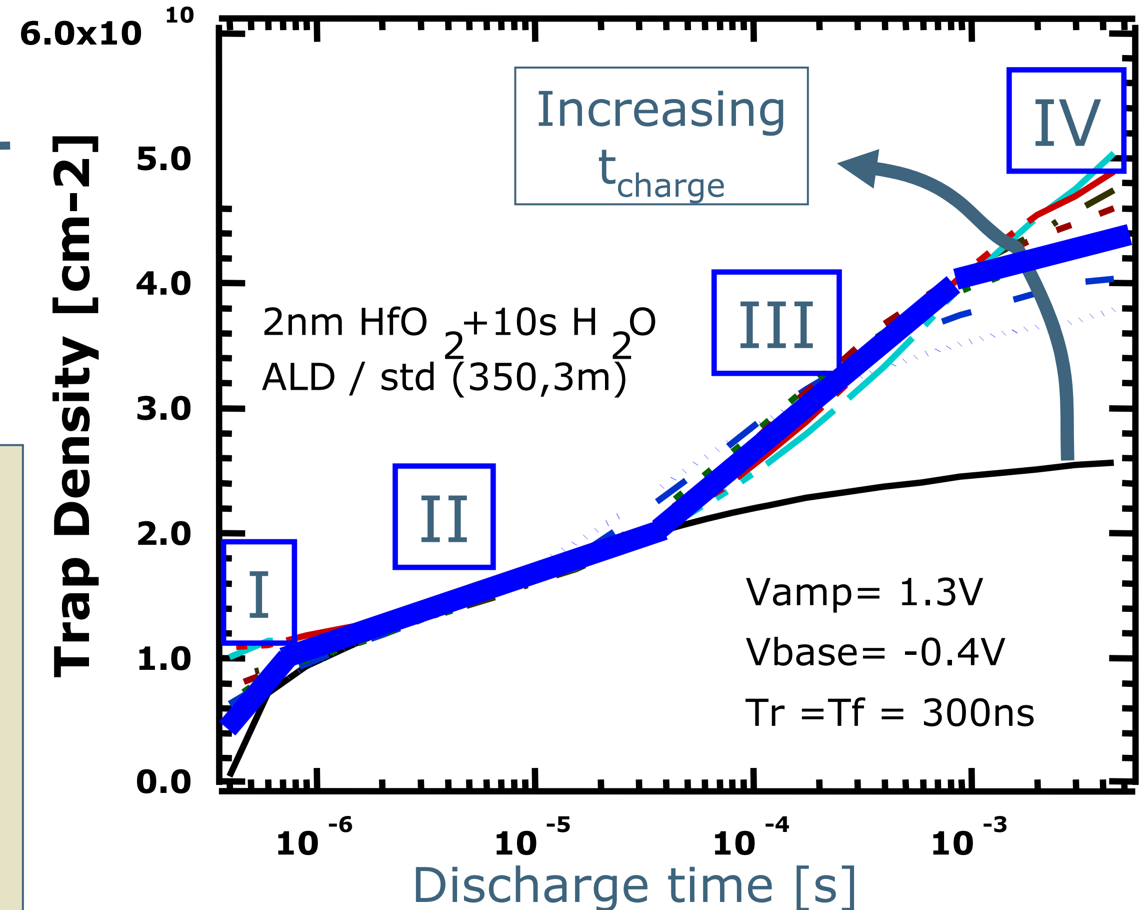
Parameters not completely separated !

Variable t_{charge}-$t_{discharge}$ charge pumping (VT^2CP)



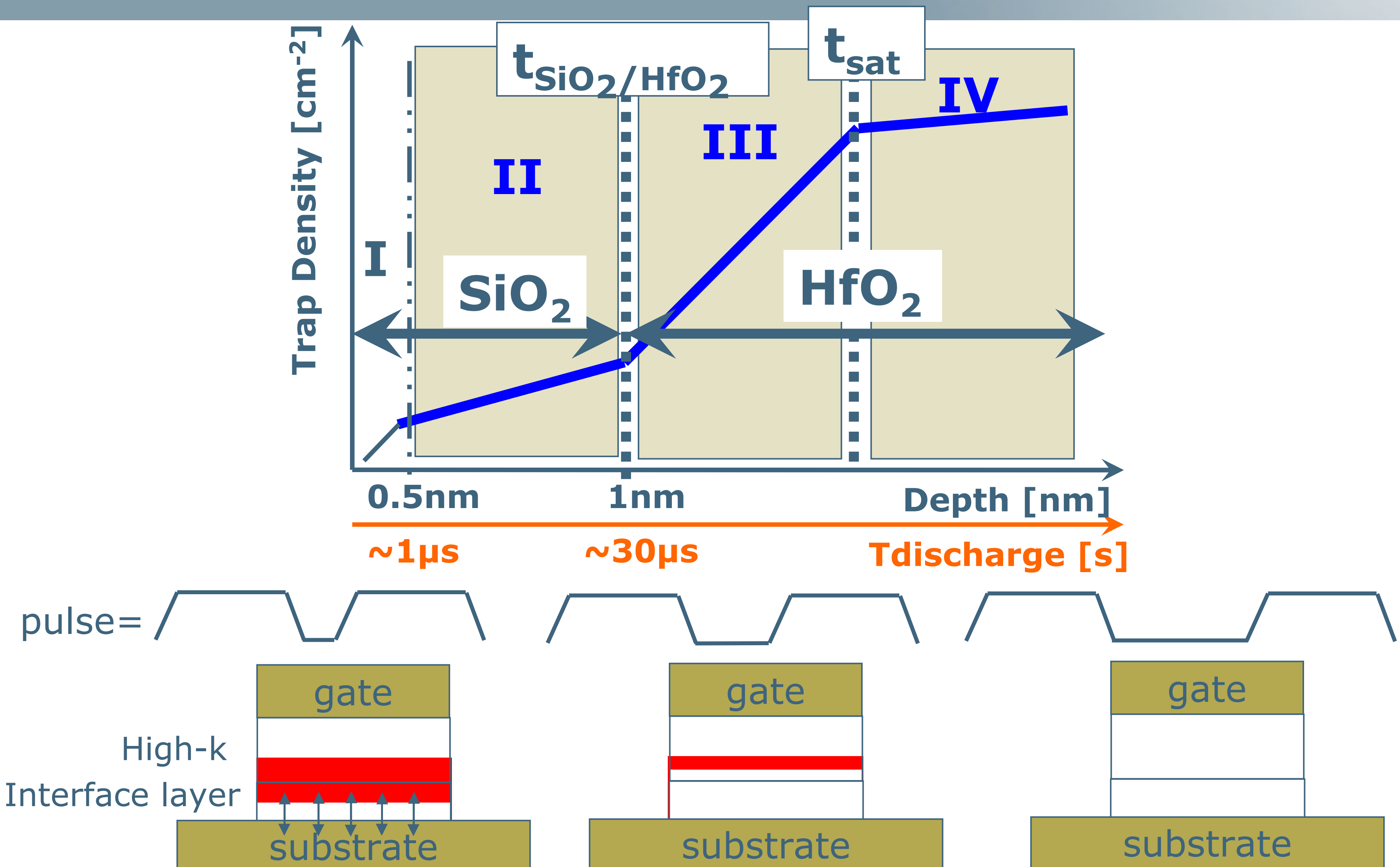
Variable $t_{discharge}$
with fixed charge time
Duty cycle = variable



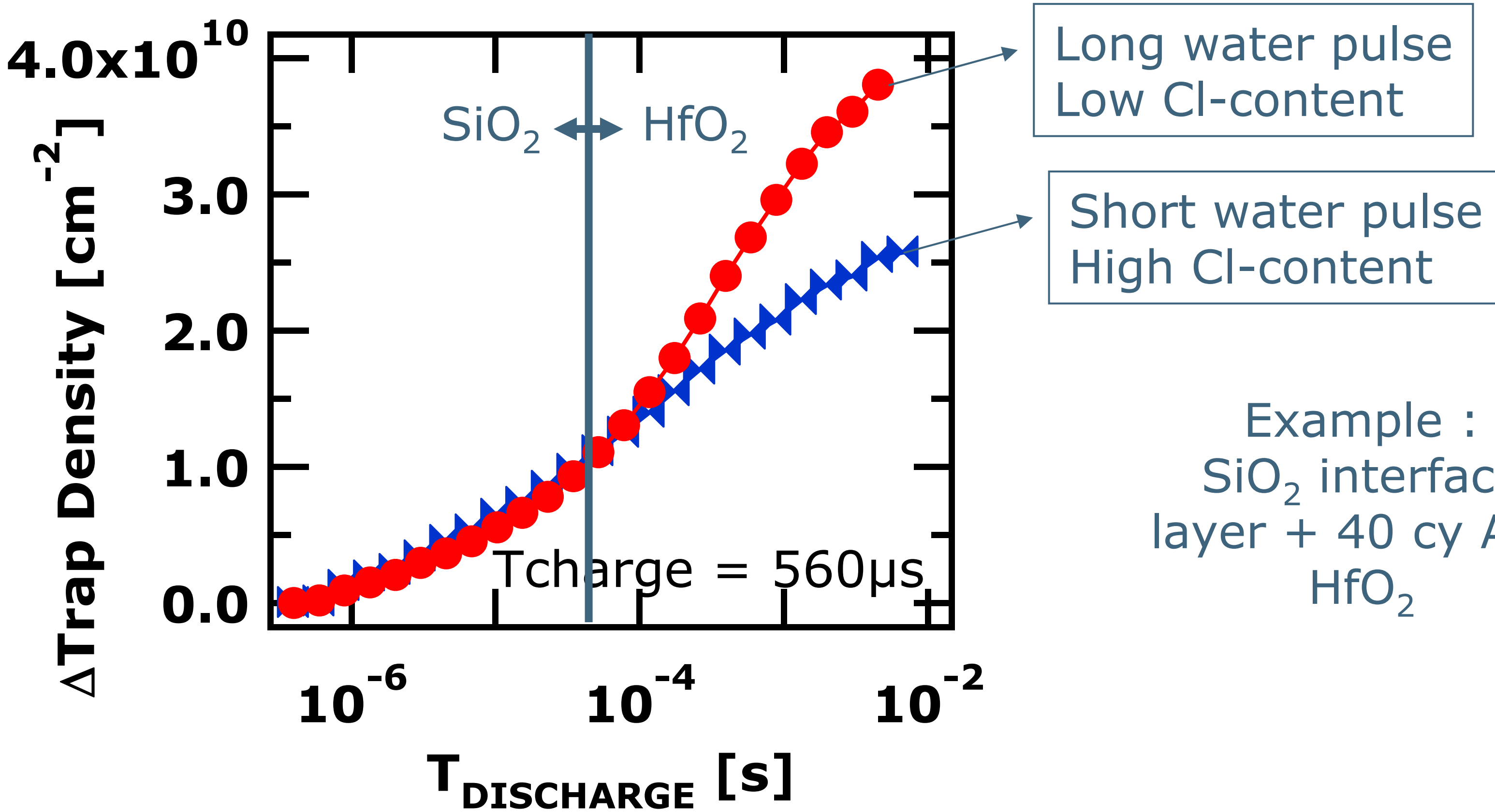


M.A. Zahid, IRPS 2007
M. Toledano-Luque, TED 55(11)

Basic interpretation VT²CP



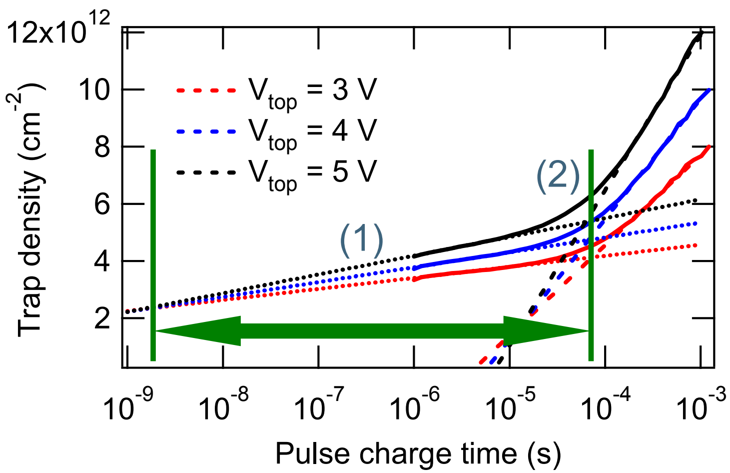
Application example 1: VT^2CP to characterize process dependence of defects in SiO_2 and high-k



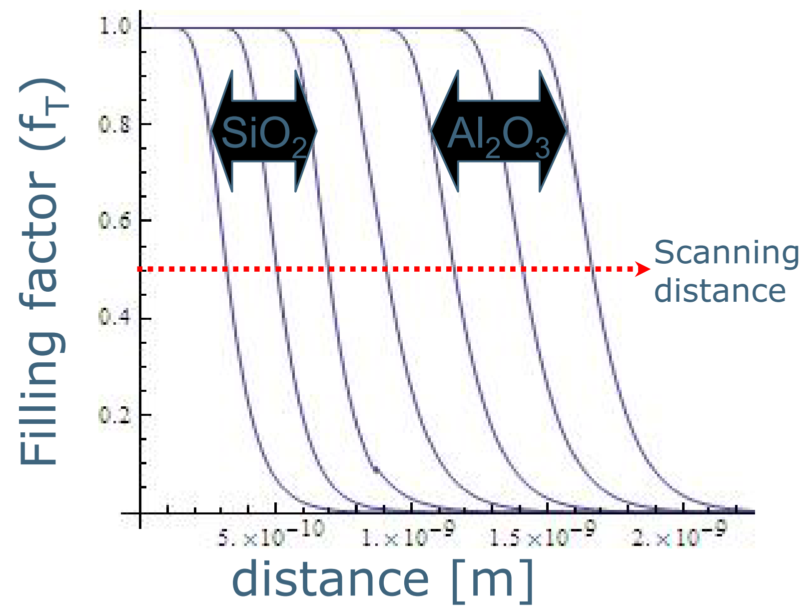
Example : SiO_2 interface layer + 40 cy ALD HfO_2

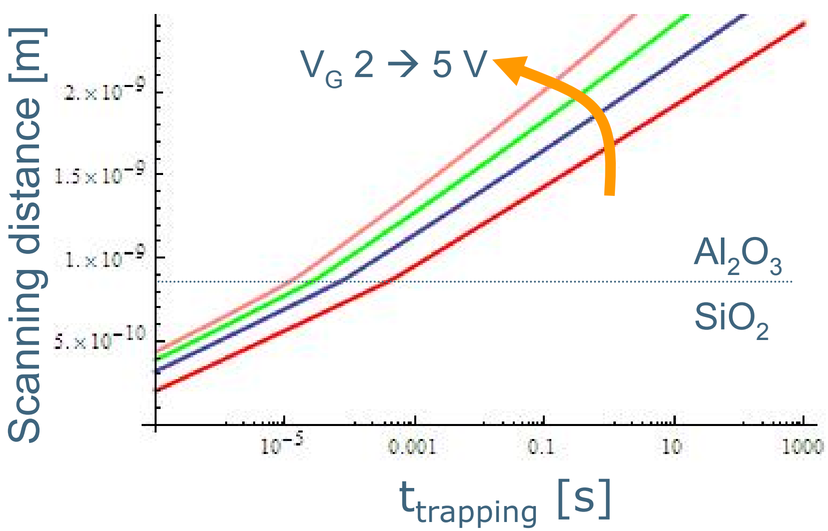
- Position information of dielectric quality

Modeling scanning distance
- Consistent with measured data



σ (SiO_2)= σ (Al_2O_3)= 2x10^{-18} [cm^2]
m*(SiO_2)=0.47m_0; m*(Al_2O_3)=0.35m_0
Ec_off(SiO_2)=3.1eV; Ec_off(Al_2O_3)=2.6eV
Si/ SiO_2 0.87nm/ Al_2O_3 10nm

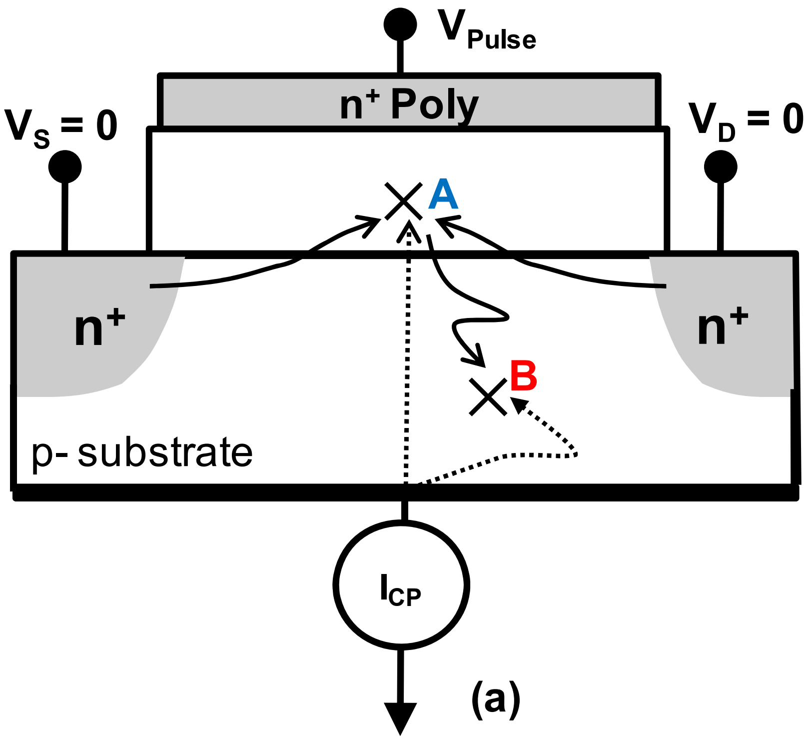


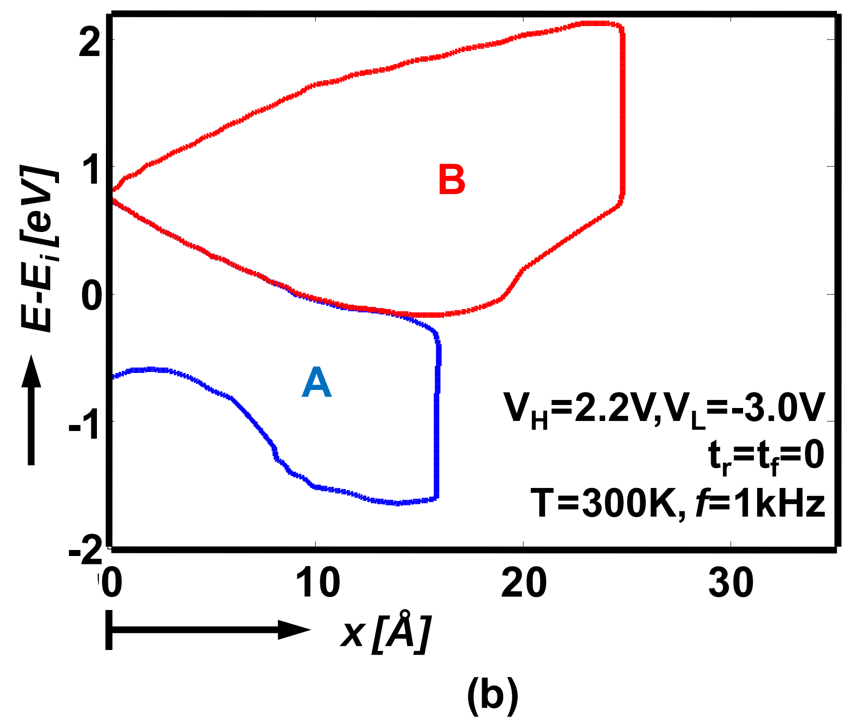


(1) Scanning rate of SiO_2 ~ 0.19nm/dec
(2) Intersection for SiO_2/*high-k*:
~60us for 0.87nm SiO_2

See Poster M.Cho, SISC 2008

Various regions accessible dependent on channel length





A = ‘classical’ charge pumping mode
B = quasi-geometric component (only for long channel lengths)

M. Masuduzzaman et al, TED Dec 2008

CONCLUSIONS

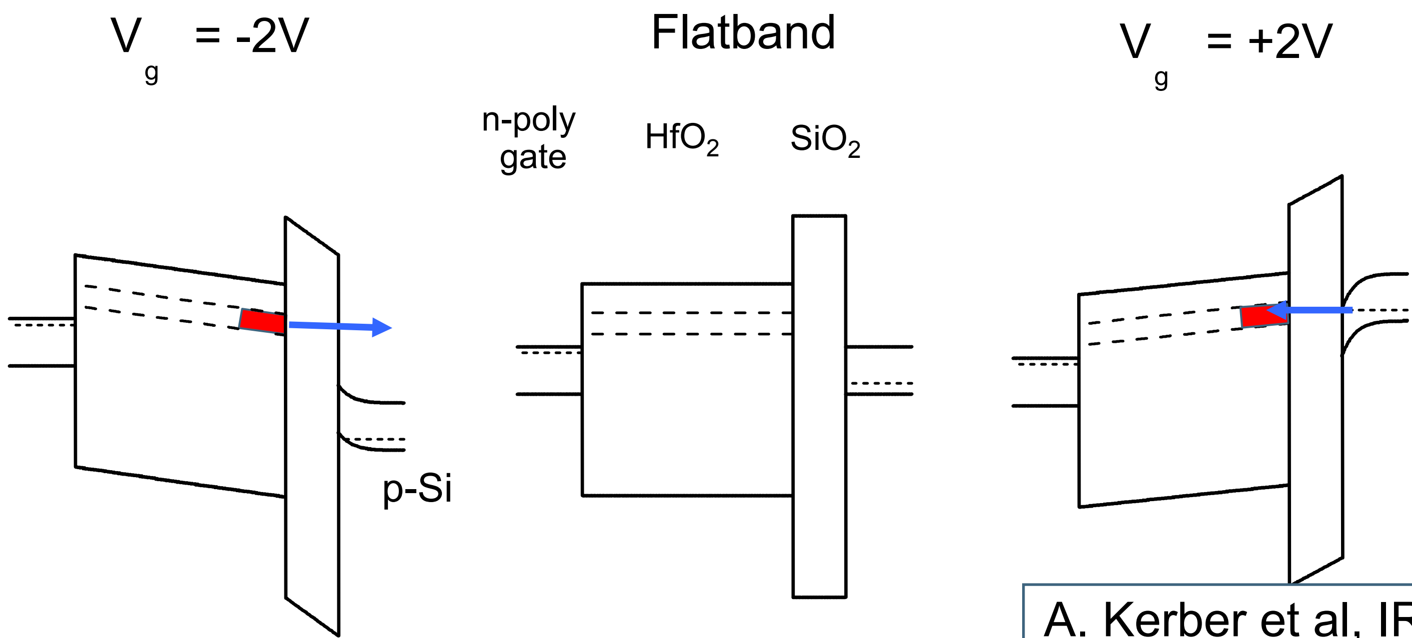
- Charge pumping is a powerful tool for MOSFET Interface characterization
- Based on a thorough insight in the physical mechanisms that are governing the charge pumping current, the interpretation of the results has been improved over the last decade, leading to a widespread use of the technique
- Charge pumping allows to determine mean values of interface trap density as well as energy distributions over a large part of the semiconductor energy gap
- Charge pumping allows to determine both uniform and non-uniform degradation damage in small area MOSFET's
- Charge pumping has proven its potentials in various fields, such as MOSFET-reliability, non-volatile memory cell characterization, SOI MOSFET characterization, radiation damage, a.s.o

Limerick #3

This is all I had this evening my dear
I hope that everything I told you was clear
And if it wasn't, I'm sorry
But let me tell you, don't worry
It will become much simpler after a beer

Thank you

Model for instabilities in high k dielectrics



A. Kerber et al, IRPS 2003

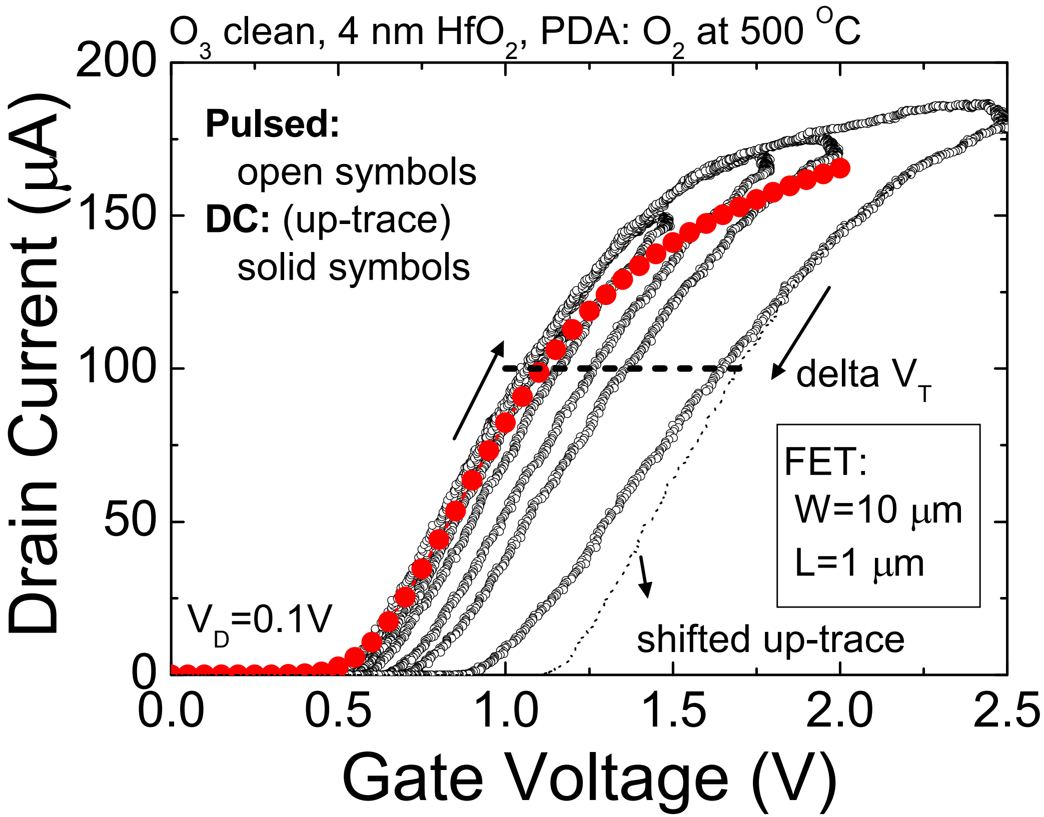
Basic features:

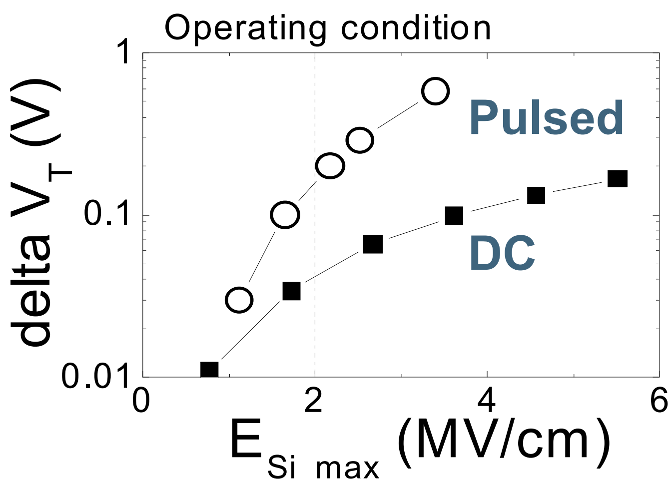
- At flatband condition defect band is located above Ec in the Si
- Defect band near the SiO2 layer moves 'fast' with Vg:

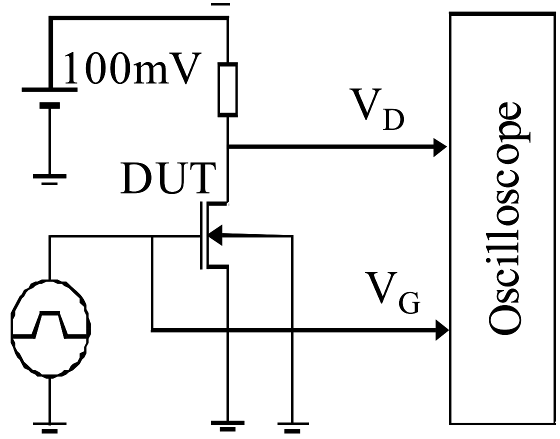
$$dE/q = dV_g \cdot (t_{SiO2}/EOTstack)$$

- Efficient charging for positive gate bias
- Efficient discharging for negative gate bias

V_t-instability in SiO_2/HfO_2 stacks:
Comparison of Pulsed and 'DC' measurements







V_t instability due to charging is underestimated by 'DC' measurements

Charging is leaking out during slow measurements

For application, pulsed measurements more relevant

A. Kerber et al, IRPS 2003